

Making Water Run Uphill

THE Annual Meeting of the British Association this year will not be the last or even the last but one, and indeed, if the programme arranged for Swansea is anything to go by, the occasion could mark a change in the association's intellectual fortunes. For a great many years, people have complained with some justice that the association's most spectacular public performance had become too dry and too academic to command the interest of younger professional scientists or even to enliven public awareness of what the association is for. Adversity seems to have brought out the best in the old organization, which for the past few years has steadily improved the sheer mechanics of presenting a kind of review of what is happening in professional science to an audience of interested bystanders and professionals. At the same time, the organization has taken, admittedly a little late in the day, some steps to put its constitution in order. That there should be talk of electing the members of the association's governing body, which has previously perpetuated itself from decade to decade, is a sign of how far the reforming spirit has gone. Yet it remains the sad truth that the sheet anchor of the association in the past few years, the annual subvention of £12,500 from the Department of Education and Science, is likely to end a year and a half from now. What will happen then? Will the British Association founder altogether? Or will it find some other way of carrying on?

The first thing to be said about the latest misfortune to befall the British Association is that it is not a misfortune at all but an opportunity. By all accounts, the Department of Education and Science has taken the view that if there is a public need of a forum for the discussion of issues in professional science or even in the relationship between professional science and society at large, then the government has itself made provision for this need in the setting up of committees such as the Council for Scientific Policy. Although there is plainly a great difference between the functions and the styles of the two bodies, there is something in this riposte. At the same time, the Department of Education seems, in saying no to the association's most recent request for money, to have declared that if public funds are to be spent on educational objectives, they had better be spent through an agency more directly under Mrs Thatcher's thumb. Galling though it may be for the British Association, and the host of voluntary workers who have over the decades helped to sustain vigorous programmes of what may be called local education, this is also a fair if ungenerous point. Indeed, what the Department of Education and Science is saying is that it would itself prefer to organize the machinery for carrying out the functions for which it is directly responsible. As a principle of public administration, that point of view is quite impeccable, for then everybody can be sure that responsibility lies squarely where it belongs. But there is also a political reason for this rigour. Why should a government trust a voluntary body such as the British Association to do its bidding—and why should a voluntary body such as the British Association assume that it will never be in conflict with the government even

on matters such as the education of the young and the enlightenment of the public?

The truth, of course, is quite the opposite. If the British Association has a place in British public life at all, it must surely stem from precisely the kind of independence that must affront politicians of all parties. At the beginning, in the early nineteenth century, the British Association came into existence because of the conviction of its founders that the institutions of the establishment, the government departments and the Royal Society included, were unable to tackle problems within the scientific profession and problems of the relationship between professional science and the wider community. In the half century that followed, the association built up a powerful record of achievement. Not only did it sponsor programmes of 'scientific research of great distinction—the Gassiot committee, one of the best known of the Royal Society's standing committees, is, for example, a legacy from the British Association—but it also served as a valuable means of awakening public interest in professional science. The principle that members of the association may be entirely unqualified as scientists is a relic of the best of nineteenth century liberalism. It is, of course, anomalous and disappointing that with the passage of decades, the involvement of the British Association with the problems which it first of all set out to tackle has if anything diminished even though the need has grown.

This is the analysis which the association should not shrink from at its present meeting in Swansea. Over the years, there has grown up a tendency to believe that everything would be different if only some sufficiently resourceful man could make impossible economies or wheedle improbable benefactions out of private industry. The job that needs doing, as much now as in the early nineteenth century, is, however, one that requires independence, both from government and from industry. In the past year or so, it is clear that a sufficiently alert organization such as the British Association could have helped to clarify questions such as the advisability of investing public money in the new accelerator at CERN (an issue which is still evenly balanced), the issue of what should be done about the organization of the research councils and the decision (scandalously delayed) on the future scale of university finance and the relationship between universities and the polytechnics. British public life, not merely British science, is impoverished by the lack of an informed and influential platform from which questions like these can be argued intelligently. Moreover, there is every sign that the scientific community would be prepared to invest not merely money but time in organizing such discussions. The fact that so many people troop devotedly and usually at their own expense to the annual meeting of the association is a sign of what wellwishers are prepared to do. To say that the British Association should become a kind of pressure group is to go too far, for these are not circumstances in which there is a single consistent point of view to be advocated. But these are circumstances in which intelligent public discussion is bound to be a social benefit, both for professional scientists



and for the community at large. To let this potential role go by default would be a loss for British public discussion as well as a tragedy for the British Association.

But how is it possible to do ambitious things without a budget? This is what the accountants are bound to say. Undoubtedly there is much that could be done to turn this criticism by strictly business stratagems. At the same time, however, it is essential that means should be found for simplifying the association's constitution. The issue is not whether the sixteen sections into which the proceedings are divided should continue to exist, but whether there is some device for allowing the organization as a whole to act decisively when its own interests and those

of its members call for something to be done. At the same time, ways must be found of convincing even busy working scientists that the association can do something to further the common cause of a healthy scientific community in an appreciative and responsive setting. In attempting all this, organization in the ordinary sense is much less important than publicity of the right kind, but everyone knows that to say as much will offend against the over-modest dissembling of the British Association's members. This is why it is to be hoped that the annual meeting in Swansea will find people prepared for the radical departures from past policy that the circumstances warrant.

The Next Step with Nuclear Explosions

THE French government has done the decent thing in bowing to public pressure and cancelling the series of nuclear explosions planned for the next few weeks in the atolls of the Tahiti group. For one thing, of course, it is intolerable that the government of France (like that of China) should continue to release large quantities of fall-out into the atmosphere without acknowledging that to do so is to create a public nuisance. This, after all, is a form of pollution that justifiably merits concern. Protests at the impending explosions from neighbouring regions in the Pacific may have helped to change the French government's course of action, but it is also probable that the time is coming when the Gaullist view of the utility of nuclear weapons is fading. It is a long time now since the strategic justification of the French independent nuclear force was argued with enthusiasm. For one thing, it has turned out to be a costly way of hoping to make sure that in the event of an aggression, it would be possible for the military forces of France "to tear an arm" off the enemy before sinking beneath a hail of nuclear explosions. Second, nuclear weapons seem increasingly irrelevant to the military problems of Europe, especially now that the gap between East and West has been bridged to some extent by the continuing civility of relations in central Europe and by the signature last week of the agreement between the Occupying Powers on the future of Berlin. Third, the French nuclear force tends to be a device within the framework of an increasingly united Europe (and the British nuclear force, now almost forgotten, would be equally an irritant if it were ever advertised). In short, the government of France would be well advised to wash its hands of nuclear weapons at this stage of the game. Enough has been done to show that the properties of nuclei are the same in France as elsewhere. Not nearly enough has been done to make use of them.

In these circumstances, the time has come rather to ask what arrangements will in future be made for regulating the uses made of nuclear weapons. It is getting on for a decade since the partial ban on nuclear tests was negotiated, and the International Atomic Energy Agency in Vienna is now up to its eyes—and some say buried—in the intricacies of policing that unsatisfactory agreement. In the year just ended, the agency took over responsibility for policing safeguards agreements at about a hundred nuclear installations. According to the annual report for the year just past, it spent a total of \$1.7 million on the

work, and the budget for the coming year, also now published, suggests that this will rise to \$2.6 million in the next twelve months. As always, the agency's abiding problem in the implementation of the safeguards policy is the recruitment of sufficiently skilled people and the determination of the conditions under which they should work. Optimism at Vienna runs as high as ever, but it is hard for outsiders easily to accept that the agency would be able to police effectively safeguards for a world in which the number of nuclear power stations under its authority was counted not in tens but in thousands.

This is why it is heartening in the past few months that the subcommittee of the UN Assembly on Disarmament, meeting at Geneva, has taken up again the issue of extending the partial test-ban to include all experimental nuclear explosions, whether underground or in the air. Not merely would this device let the Vienna agency escape from the invidious position of policing the unpoliceable into which it has been jockeyed, but it would solve a host of other problems, from the proper balancing of risks and benefits in some of the newer developments of military technology—multiple nuclear warheads, for example. Moreover, an extension of the test-ban would help to ensure that the asymmetries incorporated in the present treaty, and which in particular allot privileged positions to countries which have already manufactured nuclear weapons for themselves, will not at some stage lead to demands among the non-nuclear powers that they too should have weapons to themselves. In this sense, the non-proliferation treaty remains what it has been since its conception (and what its preamble says it is)—a temporary expedient.

As it happens, technical developments are also favourable. The article on page 51 by Dr F. M. Anglin shows how it is now possible to discriminate more accurately than previously between earthquakes and nuclear explosions with less than the equivalent of 10,000 kilotons of energy. The idea (see page 8) is that even quite small explosions can be distinguished from earthquakes of comparable size by the high frequency parts of their seismic spectra. To be sure, nobody would pretend at this stage that all explosions can be distinguished unambiguously from naturally occurring earthquakes, but there has been an immense improvement in the technology of test detection and discrimination since the days when the non-proliferation treaty was first advanced.

Making Voters of Universities

THE United States Senate has done as much as any legislative body can to endear itself to academics. It remains to be seen whether it will be able to persuade other politicians in the United States, and especially the Administration, that the time has come to be generous once again to colleges and universities in the United States. Two weeks ago, more or less with acclamation, the Senate passed the draft of a piece of legislation intended to replace the Higher Education Act of 1965, and which would have the effect, for the first time, of making attendance at an American university into an assurance of a maintenance grant, at least for students whose parents are not too well off. But as well as helping needy students, the proposed legislation would help needy universities and colleges—those on the institutional breadline would be helped not merely with matching grants to accompany the student grant but also with special awards in times of trouble (see *Nature*, **232**, 518; 1971).

By backing this bill, the Senate has done three important things. First, for the first time, it has declared that higher education in the United States is something which the federal government must worry about. So far, apart from some of President Johnson's devices for helping universities in need by providing funds for libraries and other special facilities, the operations of the National Science Foundation and the other research bodies which for practical purposes support graduate education but only fitfully, and the scheme for guaranteed student loans introduced a year ago and still too new to be assessed, the United States Administration has traditionally taken the view that good colleges and universities will be able to swim and the others deservedly to sink. It will be interesting whether the Administration, Republican as it turns out, will accept an argument that has failed in previous Congresses to enchant its predecessors.

The Senate has also affronted a great many of those who believe that higher education, like the operation of supermarkets, should find its own level by proposing that a large amount of money should be spent on subsidies. On present estimates, the new higher education bill, if it ever becomes law by the summer of 1973 (when the present Act runs out), would provide \$18,000 million over three years. Many of those who have accepted the principle that higher education should properly be a charge on the federal government have found their hands rising in horror at the prospect that the rate of public expenditure on universities and colleges should rise to such a point that it exceeds even the spending of the National Aeronautics and Space Administration and other similar agencies. But, of course, this is only one side of the coin. The better half of the case is with those who argue that if a government such as that of the United States can spend \$5,000 million (at the peak in 1968) on getting to the Moon and related activities, surely it can find a comparable amount to ensure the intellectual future of American society and industry.

That, of course, is the rhetoric. The interest of what the Senate has done, so soon after a period in which universities of all kinds and research academics in particular have been acutely short of funds, is that it presents

the Administration with an important political challenge. A year from now, when the Senate's bill should have found its way through yet another round of committees and should have struck a response from the House of Representatives as well (where similar ideas are waiting in the wings), President Nixon will be up to his neck in the election that will decide whether or not he will spend a second term at the White House. Although there is no strong tradition, in the United States or anywhere else, that higher education can become a political issue, next year may provide an exception. After all, the past few years have done very little to heal the breach between the present Administration and the universities which was created as much by brashness and bad manners as by malevolence in the early days of Mr Nixon's Administration. It is improbable that Mr Nixon will now willingly take the risk of giving offence in the universities even if he manages to return from Peking next spring with quite unexpected political triumphs. In all the circumstances, this may be an occasion when the Administration finds it prudent to let the Senate have its present bill.

100 Years Ago



PARIS

Académie des Sciences, August 21.—M. Faye in the chair.—M. Chabris has calculated the quantity of nitrate of ammonia yearly carried down to the soil for the nutrition of plants by means of the rain to be two pounds of nitric acid, and consequently, three pounds of nitrate of ammonia per acre.—M. Dumas contributed to the Academy a piece of bread, the provision for the army, which had been infected by *Oidium aurantiacum*. Such facts are not exceptional, principally in very hot weather, and may be detrimental to the public health, as the fungus spreads very rapidly, and it is very difficult to get rid of it. A special committee has been appointed to prevent the infection if possible. MM. Dumas, Baron Larey, Tulasne the botanist, and Pasteur, the celebrated author of so many works on spontaneous generation, are members.—M. Berthelot has examined most carefully a piece of carbon from the Cranbourne meteorite, an Australian stone, and shows by many scientific arguments that the Cranbourne carbon is quite unlike the Orgueil meteorite carbon (a French specimen). The Cranbourne carbon must have been acted upon by a high temperature in ultra-terrestrial space, and no trace of organic origin is to be found on it.—M. W. de Fonvielle sent through M. Leverrier a note establishing that meteoric phenomena analogous to the Marseilles phenomenon are not exceptional cases.—M. Bert, the former Prefect of Lille during the war, described some most interesting experiments on the effects of pressure in suffocating animals living in a confined space. The rapidity of death is not the same for every kind of animal. If the pressure is very high the death is not due to any mechanical effect or to the want of oxygen, but to the presence of carbonic acid, resulting from respiration. It is poisoned by the produce of its own lungs.—The Academy held a secret committee for the nomination of a free member. The list of candidates long delayed was at last published, and M. Belgrand is at the head. But the nomination will be contested.

From *Nature*, **4**, 380, September 7, 1871

OLD WORLD

POLLUTION

Swansea Syndrome

A GRANT of £11,000 has been awarded to Mr G. T. Goodman of the Department of Botany at the University College of Wales, Swansea, so as to continue his investigation of non-ferrous metals in mosses, grasses and soil. A first report of this study, in the region around Swansea, appeared in June this year (Goodman, G. T., and Roberts, T. M., *Nature*, **231**, 237; 1971). Now the Natural Environment Research Council is prepared to support an extension of the study to the rest of Britain.

The original discovery showed the presence of large amounts of metals such as zinc, lead, cadmium and copper in plants and soil to the east of the Swansea Valley. At one of the most heavily contaminated sites, for example, some mosses contained as much as 1,200 parts per million of zinc, 800 parts per million of lead and 40 parts per million of nickel. One bizarre accompaniment of the investigation was the discovery that several horses found dead in a field near Llansamlet in the spring of 1970 had large concentrations of lead in their systems, although it has not been possible to show beyond dispute that lead poisoning was responsible. Malnutrition is another possibility.

The appearance of the article by Goodman and Roberts has led to several decisions, not the least important of which is the setting up by the Welsh Office of a working party on pollution by metallic elements in and around the Swansea Valley. It is assumed that the high concentrations of metal elements in the region have been caused by windborne pollution from neighbouring slagheaps, but another possibility is that exhausts from chimneys are responsible.

The working party is interdisciplinary in outlook and the scientists under the chairmanship of Dr R. T. Bevan, Chief Medical Officer at the Welsh Office, include Dr D. H. Peirson of the Atomic Energy Research Establishment, Harwell, Dr S. R. Craxford of the Warren Springs Laboratory, Professor C. R. Lowe and Dr J. R. Glover of the Welsh National School of Medicine, Professor P. J. Lawther of the Medical Research Council and Mr G. H. Francis of the Agricultural Development and Advisory Service. The first meeting of the working party was held in July and the second is scheduled for next week. The brief given to the working party was to recommend to the Secretary of State for Wales the sort of research programme needed to be set up to extend the work of Goodman and

Roberts with particular reference to any correlation that might be established between the presence of heavy metals in the environment and human and animal life.

A member of the working party said this week that several botanical sampling stations would be set up throughout the area and that direct air sampling would also be carried out. Soil and herbage samples would be analysed and blood and urine samples would be taken from hundreds of people throughout the area. It is planned that this survey will take twelve months to complete.

The simplicity and effectiveness of the technique developed by Goodman and collaborators make it particularly suitable for determining heavy metal contamination over even a short period. In areas that are moss deserts—because of high pollution—it is possible to deposit moss samples for as little as two weeks and obtain information on the accumulated metals during this time. The inexpensive nature of this research—the original investigation was carried out on a £500 budget—makes it a particularly useful tool to monitor atmospheric pollution. The problems in the Swansea area, however, might turn out to be insignificant compared with those found elsewhere in the country but there is no doubt that the working party will highlight the effects and dangers of a high atmospheric heavy metal concentration.

EMPLOYMENT

Graduate Problems

FIRST reports from the universities suggest that the plight of applied scientists may not be as bad as was predicted earlier this year. But times are still hard for pure scientists. In April the Confederation of British Industry (CBI) predicted a cut of between 15 and 20 per cent in the largest traditional field of graduate employment—industry and commerce. At the same time, several companies cancelled their visits on the traditional “milk-round” to British universities.

Now that the degrees examinations are over, what is the situation? The university appointments boards emphasize that it may be too soon to know for sure, partly because some graduates are on holiday and not bothering to look for jobs and partly because they often do not let their appointments boards know after they have accepted a job. But the impression is that it has been a difficult year. Quite how difficult depends on where you look. At Oxford, Mr T. Snow, the Secretary of the Appointments Committee, said that “graduate unemployment is not a problem”. But he does admit that there

are fewer jobs, particularly for those who want specific careers. It is particularly hard for chemists who want to remain chemists.

The estimate of the size of the problem varies from university to university—at Sussex, for example, Mr I. H. F. Kerr, the Appointments Officer, described the situation as “very serious” and estimates that as many as 400 of the 900 ex-students may still be job-hunting. At East Anglia, the number still looking is “higher than last year, but not dramatically so”, while at the University of Manchester, figures for mid-July show that the applied scientists—with the notable and unsurprising exception of aero-engineering—were managing as well as or only slightly worse than last year, but that the pure scientists are finding it very bad. Half of the physicists are unemployed, however, compared with a quarter last year, and two-thirds of the chemists compared with half last year.

Yet things are not disastrous and there are hopeful reasons. The employers' concept of a “graduate job” is beginning to change. Companies are considering graduates for jobs that people with A-levels would have had a few years ago and they are also more ready to accept applicants with any degree, not just in a specific field. Graduates are equally beginning to look at careers that formerly they would not have considered—chartered accountancy, insurance, banking and local authority work. And for the graduate who is prepared to be flexible there are still jobs available. The University of Manchester's latest Clearing House List has some 1,400 vacancies (although this is a drop of between one third and one quarter on last year). Many science students are turning to completely different fields—British Rail travel planning or the Civil Service, for example. The theory that a university degree is a passport to a certain sort of job seems now to be a dead duck. Today there are even graduate shift foremen in a steel firm near Cardiff.

But if the employment situation is difficult, what are the graduates doing? Many are still on holiday or doing temporary work, and will only start looking in the autumn when the situation may well be worse, not better. Others may be trying to postpone the decision by prolonging their education. There has been an increase of more than twenty per cent in graduate applications to teacher training colleges—nearly 14,000 have applied for an estimated 9,000 places. (The new requirement that teachers in secondary schools should have a Dip.Ed. may help to account for this.) Some appointments officers also suspect that more graduates are simply dropping out, saying that “no job is better than any job”.

INDUSTRIAL RELATIONS

Jack Hill

THE dismissal of Mr Jack Hill's case against his employers highlights the predicament of up to 60,000 engineers throughout the country. Mr Hill, a sixty-three-year-old contracts engineer with the Newcastle upon Tyne engineering company of C. A. Parsons, was given a month's notice at the end of July for refusing to join the Draughtsmen's and Allied Technicians' Association (DATA) which, after a three-year battle, had won the right to a closed shop at Parsons. He sued the firm for wrongful dismissal.

Mr Justice Brightman, before whom the case was argued, concluded that under the present law he could not prevent the company from dismissing Mr Hill. He added that if the law had allowed, he would have granted the order. Under the Industrial Relations Act, which comes into force by stages from October, the action of C. A. Parsons would be illegal. Mr Hill is appealing against the decision, but it is likely that the appeal will be heard before the relevant clause of the new Act comes into effect. Until the appeal is heard Mr Hill retains his job.

There are thirty-eight engineers at Parsons in the same position as Mr Hill, but the United Kingdom Association of Professional Engineers (UKAPE), of which Mr Hill is a member, is only too aware that as many as 60,000 engineers in 5,000 companies around the country are in a potentially similar position. They are all a minority body of professionals in companies with large numbers of technicians belonging to unions such as DATA.

UKAPE sees itself as protecting the professional ethics and integrity of senior engineers and feels that some of the aims of the Amalgamated Union of Engineering Workers, of which DATA is a part, are directly opposed to that integrity. As things are, UKAPE seems to be able to say plenty but to do little—its membership is not all that large. But if those whom UKAPE considered as professional engineers—those with degree or equivalent status—are to protect what they consider to be their rights, then once the Industrial Relations Act comes into force they will have to join UKAPE if they are to have any say in negotiations between management and the proposed agency shops. Otherwise they will find themselves out in the cold, a minority without national support among a hostile and militant trade union. In such a situation, UKAPE could become a force to be reckoned with.

The circumstances of professional engineers are by no means unique. Chemists and physicists working in industry could find themselves in the

same position. So far, to be sure, they have not had to deal with unions as militant as DATA, the members of which worked to rule twice and struck once in their efforts to establish a closed shop at Parsons.

An Association of Professional Scientists and Technologists is being formed as a trade union (like UKAPE) although it will not be affiliated to the TUC (also like UKAPE) so that it can protect its members in the agency shops. But if Mr Jack Hill stands as one of the last—if one of the most spectacular—victims of the present law, it may transpire that his case is only the opening shot in a battle to come, for although the Industrial Relations Act will prevent closed shops in their present form, there are some fears that militancy will find other ways of achieving the same result.

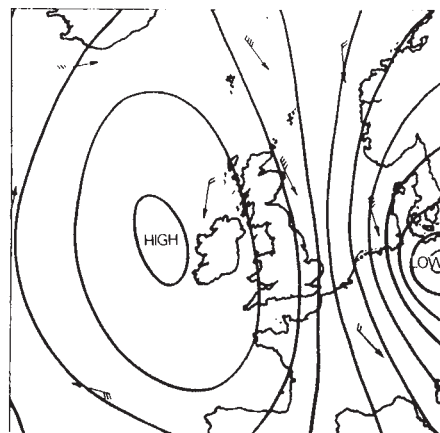
RIVER FLOODING

London Sits and Waits

LAST week, five British Government departments issued a joint circular spelling out the consequences of the flooding of the Thames. It recommended all organizations likely to be affected to review their precautions and outlined the warning system at present in operation. London seems still to be in danger seventeen years after the Waverley Committee urgently recommended that it should be protected. Three hundred people died the year before, when the Thames Estuary flooded.

The Waverley Committee had good reason to be concerned. If the present level of defence was overtopped by 1 foot, serious flooding would result, with damage likely to be in the order of £1,000 million quite apart from the potential loss of life and the inconvenience of having London partially paralysed for up to nine months. The area theoretically in danger in the worst predictable combination of circumstances is that below the 23 foot contour. This is an area of 60 square miles and 1.2 million people, with 15 power stations, 4 sewage plants, 23 storm-water pumping stations, 56 telecommunications buildings, 46 miles of underground and 70 stations, almost all Government departments and the Houses of Parliament. The problem arises first because South-East England is settling so that the exceptional high water mark is rising at the rate of about 2.43 feet a century and second because of the power of the North Sea surges, usually caused by an intense depression from the Atlantic closely followed by a high pressure area. If the low pressure trough, on reaching North-East Scotland, swings south-east over the North

Sea, the surge, amplified by the continental shelf, the comparatively narrow channel and strong winds, rushes down the North Sea and into the first convenient opening—the Thames. Every depression causes a surge, although usually a minor one. Should such a happening coincide with a high spring tide, and if the river were also full after heavy rain, the results could be disastrous.



Pressure conditions that will tend to induce Thames flooding.

At present there is a complicated five-stage system giving warnings ranging from a 12-hour-preliminary to a one-hour-final in Central London. The preliminary warning is given between 20 and 25 times a season, the three-hour Central London warning has been given only three times in the past five years. Given the present circumstances, the system seems as satisfactory as possible.

The government is at present paying 65 per cent (the Greater London Council is paying the other 35 per cent) of the £5.5 million needed to raise the riverside walls eighteen inches—the first time they have been raised since 1930. This will be completed in just over a year—the odds that the present level will be reached are about 1 in 10 in any year. With the new walls built, the odds of there being six inches of flood water in any one year will be 1 in 110. To provide what the GLC describes as an acceptable level of protection—7,000 to 1 in any year or 100 to 1 in a man's lifetime—the walls would have to be raised 6 feet, which is clearly unacceptable.

The alternatives—to build a barrage or a barrier—have been bandied about various committees and engineering companies since 1954. The government has recently approved the construction of such a barrier, which is now being designed, at Silvertown in the Woolwich Reach. Design is expected to take another eighteen months to two years, and construction a further five, at a cost of about £50 million.

NEW WORLD

Patent Policies Aim to Provide Incentives to Industry

by our Washington Correspondent

A POLICY designed to enhance the likelihood that research financed by the taxpayer will be commercially exploited was announced last week by the Office of Science and Technology. The new policy should delight private corporations, which now have the chance, under certain circumstances, to gain exclusive licences for exploiting government inventions, and it should also clear up a few vexing irregularities in the Federal government's attitude towards the granting of patents on inventions arising out of government sponsored research. Described by Dr Edward E. David, the President's Science Adviser, as a "new method of bringing flexibility to the market place", the new policy is the fruit of a study of federal patent policies carried out under the auspices of the OST by the Committee on Government Patent Policy.

The new policy, which replaces that spelt out by President Kennedy in 1963, contains the following major provisions:

- Heads of departments or agencies of the Federal government can in certain circumstances award exclusive licences to contractors, allowing them to offset risk capital in bringing an invention to the market place. The chief criterion for awarding such a licence is simply that commercialization of an invention arising from government-sponsored research "will best serve the public interest", and that the invention is unlikely to be exploited unless the company is given sufficient financial incentive.

Built into the new policy is also the provision that licences can be awarded for inventions that are the primary object of a contract. The previous patent policy had specifically prevented a contractor from gaining greater rights than a non-exclusive licence to such inventions, and the Federal Council for Science and Technology found that some agencies had misinterpreted this clause so that, for example, the Department of Health, Education and Welfare might be prevented from granting greater rights to a health invention.

Explaining the rationale behind the decision to enable departments and agencies to grant exclusive licences for the exploitation of government inventions, Mr. O. A. Neumann, executive secretary of the Committee on Government Patent Policy, said that studies have brought to light cases where inventions would not reach the market place unless a period of exclusivity was provided to the developer to recoup his

private investment. Mr Neumann declined, however, to give details of such cases.

- If the Administration does decide to turn over principal or exclusive rights to an invention to a contractor, a licence to make, use or sell the invention will still be retained by the government, unless the head of the department involved determines that it would not be in the best public interest to do so.

- If the government itself decides to retain the principal or exclusive rights to an invention, it may also grant the contractor the right to a non-exclusive licence to manufacture and market the invention. But even in this case, the company's licence could later be revoked and replaced by an exclusive licence if the government decided that extra encouragement was needed to develop and market the invention.

- Finally, the new patent policy requires every department and agency to keep a close record of its licences, so that the effects of the policy can later be evaluated by the government Committee on Patent Policy. Pointing out that past government patent policies have been difficult to evaluate because of inadequately kept records of non-exclusive licences, Mr Neumann said last week "there is really no way for us to know what sort of activity we shall be getting into, and it is very difficult to see what the level of activity will eventually be".

Government departments and agencies are also divided over the question of what effect the new policy will have on their activities. NASA and the Department of Health, Education and Welfare, for example, point out that they have both in the past had the power to grant exclusive licences in exceptional circumstances, but each has only issued one or two such licences in the past few years. It seems likely, therefore, that departments and agencies will continue to allow commercial exploitation of inventions developed under government contract on a free-for-all basis as, for example, the Department of Agriculture did when a process to manufacture penicillin on a large scale was discovered under contract from the government. Exclusive rights are likely to be granted to companies only for inventions which would require large risk capital to bring them to the market place or where the market potential would be an insufficient incentive for the company to put up the development capital.

What the policy statement has accomplished, however, is to provide an explicit basis for the negotiation of government contracts which together account for some \$15,000 million of federal money each year. About half of this expenditure is accounted for by the Department of Defense, which has in the past consistently assigned patent rights to the contractor when the contract was first negotiated, before any discoveries had been made. This practice has, however, come in for some attack in the past for giving away public property. What those critics will make of the new patent policy should be interesting.

DERMATOLOGY

FDA in Rash Again

by our Washington Correspondent

THE Food and Drug Administration may soon seek to extricate itself from a grave dilemma by allowing a highly toxic drug to be legally marketed for use against severe cases of psoriasis, a debilitating skin disease. The dilemma that the FDA faces is that whatever action it takes on the drug will result in strong criticism; if it allows the drug to be used against psoriasis, it will be assailed for not protecting the public against a drug that has already been shown to be highly toxic, but if it takes action against the manufacturer to prevent the drug from being sold as a dermatological preparation, it will be accused of taking out of the hands of physicians their only effective weapon against psoriasis.

The FDA has only itself to blame for its dilemma, for during the past four years it has turned a blind eye on use of the drug by dermatologists, and has taken no action against the manufacturer, Lederle Laboratories, for commercially promoting the drug for an unlicensed use. It has been fully aware of the drug's severe side effects, but has taken few steps to warn either the public or the medical profession about them.

Under the trade name 'Methotrexate' (MTX), the drug was licensed by the FDA in 1959 for use against leukaemia, but in 1963 the manufacturer notified the FDA that it was sponsoring experiments using the drug to treat psoriasis. The FDA has thus allowed the drug to have an experimental status

for eight years, thereby violating its own rule that such situations should not be unduly prolonged.

Lederle Laboratories has notified the FDA of 71 different adverse side effects associated with use of MTX, ranging from disease of the kidneys, liver and blood, and ulceration, to nausea and bleeding from the nose and mouth, and the FDA is also aware that there have been at least 25 deaths associated with use of MTX to treat psoriasis. At the same time, however, some dermatologists have reported complete clearing of psoriasis symptoms from fifty per cent of patients who had the disease over as much as three-quarters of their body. Other treatments for the disease, which manifests itself as red, scaly patches which can be disabling if they cover a large area of the body, are largely ineffective.

Why cannot the FDA continue to ignore the problem as it has done for the past eight years? The chief reason is that the House Intergovernmental Relations Subcommittee which, under the chairmanship of L.H. Fountain oversees the work of the FDA, last month took the unusual step of holding public hearings on the FDA's handling of the drug. As a result, Mr Fountain has accused the agency of being "grossly remiss in not formulating and enunciating a firm but reasonable policy of enforcement", and last week he wrote to Charles C. Edwards, Commissioner of the FDA, urging him to write individual letters to physicians warning them of the hazards associated with MTX.

Fountain's letter, which was also signed by Mrs Florence Dwyer, ranking minority member on the committee, pointed out that the FDA has taken off the market drugs that are much safer than MTX, and that it would be inappropriate for the FDA to allow the drug to be marketed for treatment of psoriasis. The chief concern is that even if the FDA insists on safeguards such as strict supervision over the use of MTX, there is no guarantee that all doctors would be able to exercise the strict control that would be required, and that, in any case, adequate and well controlled studies of the efficacy and safety of the drug have not been carried out by the agency's scientists during the eight years that MTX has been on experimental trial. It is therefore argued that the FDA should warn every doctor by individual letter of the danger of using MTX to treat psoriasis, and that doctors should also be warned that continued use of the drug for unlicensed purposes could result in MTX being withdrawn from the market.

James D. Grant, Deputy Commissioner of the FDA, said in testimony before Fountain's subcommittee last month that the agency hopes soon to work out with the manufacturer an

acceptable labelling for MTX which would permit its use only in severe cases of psoriasis. Such a compromise would in effect give legal sanction to the practice established during the past four years—a practice which Fountain said left him wondering whether the FDA understands its role to "enforce the law, not to debate, not to enter into agreements, not to anguish about violations but to act and terminate violations". He added that "if the FDA will not enforce the law, then the public does not need an FDA".

ALASKA

Pipeline Economics

WHILE the Department of the Interior is busy revising its environmental impact statement on the proposal to construct an oil pipeline across the Alaskan wilderness, opposition to the plan is continuing unabated. The latest attack on the plan has come not from environmentalists, however, but from nineteen congressmen (eighteen Democratic and one Republican) who are concerned that the pipeline would exacerbate the disparity in oil prices that now exists between the west coast on the one hand and the mid-west and east on the other. Led by Mr Les Aspin from Wisconsin, a noted opponent of the trans-Alaska pipeline plan, the congressmen (who all represent midwestern or eastern states) have written a letter to President Nixon urging him to instruct Rogers Morton, Secretary of the Interior, not to issue permits for construction of the line until a full economic and environmental study of an alternative route through Canada has been carried out.

The nub of their argument is that present plans to pipe oil from the North Slope across Alaska to the ice-free port of Valdez, and then to ship it to the west coast for consumption, will decrease oil prices in the west and further increase the difference in oil prices between the west coast and the rest of the United States. A barrel of oil now costs 60c more on the east coast than on the west, and the letter to President Nixon suggests that the difference would be increased to about 140c if the pipeline were built. If oil were piped through Canada directly to the midwest and the east coast, however, the congressmen contend that the price differential would be wiped out. There is, however, no overriding ecological argument for re-routing the pipeline across Canada, although danger from spillage would be reduced since the Canadian route would not involve shipping, and unlike the trans-Alaska route, it is not located in an earthquake zone. The Canadian route would be about twice as long as the Alaskan route, it would cost consider-

ably more, and it would take two or three years longer to build. In attempting to get the pipeline re-routed, its opponents, like the opponents of the SST, will therefore face considerable opposition from industrial interests. Among the signatories to the letter is Senator William Proxmire.

Short Notes

HEW Empire Attacked

Mr Paul G. Rogers, chairman of the House subcommittee on Public Health and Environment, hopes soon to introduce legislation designed to dismantle the unwieldy bureaucracy of the Department of Health, Education and Welfare. Speaking at a meeting of the American Hospital Association last week, Rogers said that health has a weak voice and leadership in the Federal Government because it has no secretary at cabinet level to argue for and coordinate the activities of the health industry, and accordingly he hopes to set up a department of health, outside the realms of HEW. Such a department, he suggested, would obviate the necessity for setting up specialist agencies such as the Conquest of Cancer Agency and the Office of Drug Abuse Control. Claiming widespread support for this proposal, Rogers pointed out that he knows of no organization that would oppose it, but even if he introduces the legislation as soon as Congress reassembles in September, it is unlikely to pass through the Congressional mill before next spring.

Professor for Life

The University of Chicago has appointed Charles L. Fefferman, this year's winner of the Prix Salem and a specialist in Fourier series, to the rank of full professor in the department of mathematics. The department has twenty-three other professors of mathematics, many of whom are better known than Fefferman, but what makes his appointment interesting is that at the age of 22, Fefferman is the youngest full professor in the history of the University of Chicago and probably of any other major university in the United States. He entered the University of Maryland at the age of 14, with the help of a special dispensation allowing him to be admitted without a high school diploma, obtained his BS degree in 1966, when he was 17, his PhD from Princeton in 1969, and was appointed assistant professor at the University of Chicago last year. Fefferman published his first paper at the age of 15, and has eight papers now in the press. His tenure of the professorship at Chicago is for life.

NEWS AND VIEWS

Detecting Underground Nuclear Explosions at Last?

IF the results of recent work in Canada are to be believed, it is possible to locate and verify explosions of more than 20 kton, fired in consolidated rocks, using seismometers in routine production. The first announcement of this achievement came at the twenty-fifth general assembly of the United Nations in December 1970, when the Canadian delegation presented a largely theoretical assessment of the feasibility capacity today for controlling a ban on underground testing of nuclear weapons. This study was based on information supplied in accordance with a resolution (2604A(XXIV)) of the general assembly in December 1969. At the Conference of the Committee on Disarmament in Geneva this summer, Canadian workers claimed, as a result of using this information in an extensive study of earthquakes and explosions, that the identification threshold for Eurasia is about 20 ktons referred to containment in consolidated rock. (To be sure, the deliberate evasion of detection is a separate question.)

Thirteen years ago this threshold of identification may well have satisfied the treaty makers, who seemed ready to negotiate an agreement on the basis of the 1958 Report of the Geneva Conference of Experts which concluded *inter alia* that some twenty to a hundred earthquakes a year would remain seismically undistinguishable from underground explosions of about 5 kton yield given a specially deployed network of control stations. Only earthquakes would be positively identified. Stations which now routinely communicate data could, the Canadian study now claims, be improved for a moderate cost such that the identification of 5–10 kton explosions would be possible with 90 per cent probability.

The criterion on which these conclusions were based is the difference between the amplitude spectra of earthquakes and explosions. The current method of measuring this difference is to apply corrections for distance, thereby normalizing the amplitudes of the P (short period) and Rayleigh (long period) waves to the magnitudes m_b and M_s respectively. Theoretical calculations presented at Beaugency in 1964 suggested that the ratio $m_b:M_s$ might be a method of identifying explosions when Carpenter (*Vesic Report*, 4410-99-X, 119; 1966) found that the amplitudes of Rayleigh waves generated by explosions should be an order of magnitude less than those of earthquakes of equivalent body wave magnitude m_b . This result has been confirmed empirically by research at Lamont (*J. Geophys. Res.*, **71**, 4333; 1966), at UKAEA Blacknest (*UKAEA Special Report*, 46; UKAEA, 1965, and AWRE 0-67/66; HMSO, 1966) and at Ottawa (*Geophys. J. Roy. Astron. Soc.*, **17**, 1; 1969). A scientific article promised by the Canadian workers may remove the remaining uncertainties in the empirical observations, and the third part of the Cambridge series (*Geophys. J. Roy. Astron. Soc.*, **18**, 233 and 353; 1969) is expected to improve the theory.

So far as the Test Ban is concerned, however, it seems that the principal issue today is the difficulty of recording Rayleigh waves from small explosions. Compressional body waves (P) from underground explosions of 1–2 ktons have been detected and located at distances up to 80°

by arrays of some twenty seismometers (AWRE 0-33/70; HMSO, 1970), but the Rayleigh waves required for identifying explosions of this size would not be recorded beyond about 10° by long period arrays of equivalent dimensions. This may be the reason why a new look is being taken at the more easily recorded body waves (P) in which interest flagged after the Novaya Zemlya underground explosions had reduced the value of the complexity discriminant (IASPEI, *Trav. Sci.*, Ser. A, Fasc. 24, 26; 1968). The Lincoln Laboratory of the Massachusetts Institute of Technology took the initiative by estimating spectral energies in selected bands of the P-wave train and found the ratio of high frequency (1.5–1.9 Hz) to low frequency (0.4–0.8 Hz) energy for earthquakes was generally lower than for explosions (IASPEI, *Trav. Sci.*, Ser. A, Fasc. 24, 25; 1968). Intuitively one might expect the seismic energy of explosions to be concentrated in the higher frequency band than for earthquakes but the discriminant worked satisfactorily in practice only for explosions in Central Asia.

On page 51 of this issue of *Nature*, Dr F. M. Anglin of the Seismology Division at the Department of Energy, Mines and Resources, Ottawa, has combined both the complexity and a modified form of the short period spectral criteria to give a satisfactory separation of earthquake and explosion samples taken from Eurasia. The methods for measuring complexity have all taken the ratio of some measure of amplitudes within the first few seconds of a short period P-wave train to the next few tens of seconds. Anglin finds the optimum condition by taking the rectified sum of the Yellowknife array elements phased to give the maximum P amplitude, and compares the first two seconds of the P signals with the following 33 seconds. In the frequency domain the principal difference from the Lincoln Laboratory procedure is to include an f_s (Hz) term over a much wider band (0.32 Hz–5.0 Hz), and to sample the first 6 seconds of the P signal. The Lincoln Laboratory used the first 20–30 seconds.

The numbers representing these two parameters for twenty-seven underground explosions in five different areas of the Soviet Union and China, and fifty-one shallow focus earthquakes within 15° of one or other of the explosion epicentres, neatly separate the two classes independently of magnitude. Explosions with complex P-wave trains, like those detected from Novaya Zemlya, are rich enough in the higher frequencies to be included in the explosion population, and the frequency spectra of earthquakes with rather simple wave trains evidently peak at lower frequencies. There are only four samples (one explosion and three earthquakes) with magnitudes less than m_b 4.5, and because this is the magnitude range in which P-wave discriminants can make a significant contribution to the Test Ban problem, there is clearly still a long way to go. (It is important that the epicentral parameters of low yield nuclear explosions should be published.) The physical basis of both the complexity and spectral differences of seismic events is not explained, so it is impossible to predict what the result of a similar test on Nevada explosions and earthquakes would be; Lincoln Laboratory failed with these events.

OCEAN FLOOR

Standard Section

from our Geomagnetism Correspondent

ALTHOUGH the third leg of the Trans-Atlantic Geotraverse was completed less than three months ago, some of the preliminary results have already been released by Dr Peter A. Rona, the project's chief scientist. The Trans-Atlantic Geotraverse (or TAG, in the inevitable mnemonic) was recommended in 1968 by the US National Academy of Sciences and the US Upper Mantle Committee as a desirable contribution to the International Decade of Ocean Exploration. Since it began in 1970 it has been under the direction of the National Oceanic and Atmospheric Administration which has put its own ship, the *Discoverer*, at the disposal of the project. According to Rona the general aim of the programme is to establish the "first standard crustal section across an ocean basin" and it certainly looks as though that will be achieved before the project is completed in 1974.

The section in question (see diagram) is a 200-mile wide, 3,500-mile long curved strip across the Atlantic from Cape Hatteras to Cap Blanc and is calculated to link points in Africa and North America which were immediately adjacent before the onset of continental drift. The western end of the traverse also links up with the Transcontinental Geophysical Survey section, the crustal section established across the United States in 1968 by means of topographic, gravitational, seismic, magnetic and rock-dating studies.

The first leg of TAG comprised three parallel tracks across the whole section. The second involved an intensive investigation of the foothills along the eastern slope of the mid-Atlantic Ridge;

and, during the third, the ocean floor was sampled by dredge and corer from the mid-Atlantic Ridge to the African continental shelf. Although the project has an impeccable scientific basis in the academic sense (one of its listed aims is "a more thorough and valuable understanding of the past and present movements of the continents and rifting of the sea floor"), the emphasis is clearly on using the basic science as a guide to the assessment of the potential oil and mineral wealth of the deep oceans; and this is reflected in the first results to be announced.

Thus Rona reports that "extremely heavy" encrustations of manganese ore (which also contains copper, nickel and cobalt) have been found on exposed rocks in regions where Earth movements have caused fractures in the sea bed; and these deposits are believed to be thicker and more extensive than the better known manganese nodules found in other areas in recent years. Also under investigation by the TAG scientists—although no actual results have so far been given—is the relation between the rifting of a continent and the formation of heavy metals such as gold and platinum. Heavy metal ores are a feature of the Red Sea, a region thought to be in the early stages of rifting; and it is thus possible that the volcanic activity accompanying the formation of the mid-Atlantic Ridge may have left heavy metal deposits in the deep Atlantic beyond the continental shelves.

But of perhaps even greater significance are the large domes which have been found in the TAG section off north-west Africa and which resemble the oil-bearing salt domes in the Gulf of Mexico. The Atlantic domes which lie in 15,000–20,000 feet of water and some of which lie at least 800 miles off the

African coast would, if their salt content is confirmed, be the first salt domes to be reported in the deep ocean and, as such, would represent the first indication of potentially significant marine oil resources beyond the continental shelves. The hypothesis which Rona and his colleagues are trying to test is that during its early development the Atlantic must have been a body of water almost completely surrounded by land and thus in which evaporation exceeded inflow. Under these conditions salinity would increase and salt would be precipitated, forming layers (and subsequently domes) into which would settle oil-forming plant and fish remains. As the Atlantic opened further, some of these deposits would be left in the deep ocean floor and not only on the continental shelf. So far, however, the validity of this hypothesis has not been confirmed.

ASTRONOMY

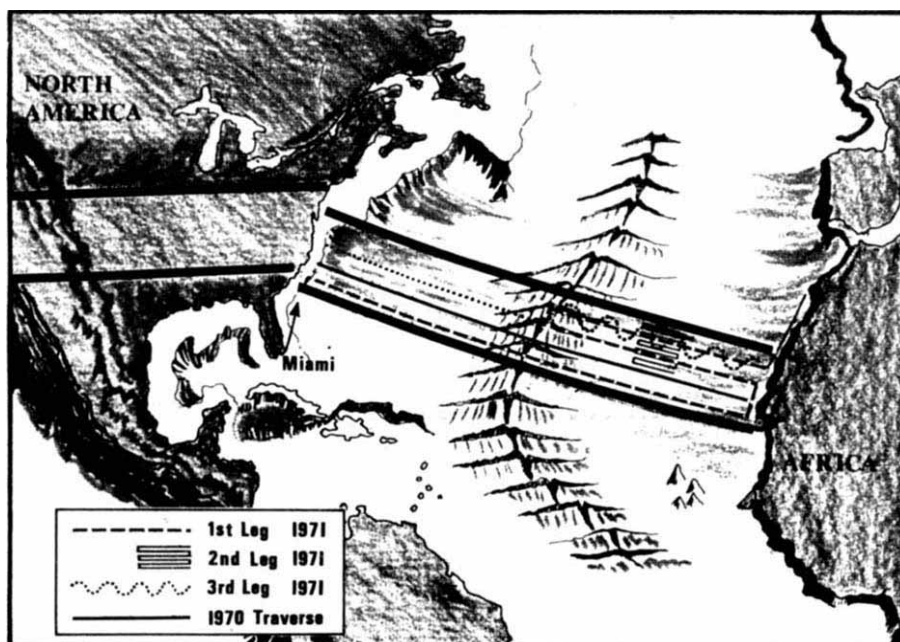
Universal Distance

from a Correspondent

How big is the universe? This was the theme of a conference on the galaxy and the distance scale held at the Royal Greenwich Observatory, Herstmonceux, from August 17 to 20 in honour of the Astronomer Royal, Sir Richard Woolley, who retires at the end of 1971.

The conference covered the whole range of distance determinations, from nearby stars to galaxies and quasars. Distances of galaxies are found by a series of intermediate steps involving galactic clusters, cepheid variables, ionized hydrogen regions and galaxies themselves as standard beacons; the latest results reported by Dr A. R. Sandage (Hale Observatories, Pasadena) give a Hubble time (time since the "big bang" if deceleration is neglected) of 18×10^9 years, exactly ten times the value estimated by Hubble in 1936, and a deceleration parameter $q_0 = 0.95 \pm 0.4$. This permits a universe that will either go on expanding indefinitely or fall back and oscillate, but probably excludes the steady-state theory ($q_0 = -1$). The actual age of the universe then comes out to be about 13×10^9 years, which compares very favourably with the age of $(10 \pm 3) \times 10^9$ years estimated for old stars in globular clusters on the basis of the theory of stellar evolution. The uniformity of the relationship between period, luminosity and colour of cepheid variables in different galaxies makes these estimates seem more firmly based than earlier ones.

Intermediate steps in the measuring process involve the distances of clusters and variable stars, corrections for absorption of light in interstellar space, the distance between the Earth and the



centre of our galaxy and the distances to our two nearest extragalactic neighbours, the Magellanic Clouds. The distances of clusters depend on fitting their "zero-age main sequence" in the temperature-luminosity diagram to that of the Hyades cluster which is well determined by the moving-cluster method and confirmed—as was shown by Dr O. J. Eggen (Mount Stromlo Observatory, Canberra)—by entirely independent measurements of the temperatures, magnitudes and trigonometric parallaxes of nearby stars resembling the Hyades in age, motions and chemical composition. The absorption of light in interstellar space immediately surrounding our solar system was shown by Dr B. Strömberg (Copenhagen University) to be almost entirely concentrated in discrete small clouds or large cloud complexes, with almost no continuous distribution of dark matter between the clouds.

Going out to distances more comparable to the size of our galaxy, Sir Richard Woolley described the determination of luminosities of pulsating variables by the Baade-Wesselink method of comparing actual changes in radius deduced from velocity variations with the relative change deduced from variations in brightness and surface temperature. In combination with the theory of pulsation, this method leads to quite low masses for many variables which in turn suggests that they could have lost mass in the course of their evolution. Dr W. Becker (Basle University) described a procedure for building up a contour map of star density in the disk and halo of our galaxy by photometry in red, green and ultraviolet bands of the spectrum which enables faint stars of different luminosity and chemical composition to be sorted out. Dr A. Toomre (Massachusetts Institute of Technology) showed that the distortion of the central plane of our galaxy could be explained by tidal action of the Magellanic Clouds provided that the Earth's distance from the centre of our galaxy does not exceed 8 kiloparsecs (compared with the round figure of 10 kiloparsecs adopted by the International Astronomical Union some years ago). But an attempt to improve galactic constants by studying proper motions was found by Dr S. V. M. Clube (Royal Greenwich Observatory) to lead to inconclusive results.

Are the redshifts of quasars of cosmological origin, indicating enormous distances, or are they an intrinsic effect due to other causes? Arguments on each side were reviewed by Dr M. Schmidt (California Institute of Technology), who showed that the hypothesis of an intrinsic origin would require a revision of present-day concepts in physics, but that the cosmological

hypothesis is not fatally in conflict with any results obtained to date.

X-RAY ASTRONOMY

Model for Sco X-1

by our Cosmology Correspondent

SINCE the discovery that Sco X-1 is a triple radio source with some features like those of a pulseless pulsar, there has been speculation that the source might indeed belong to the same family as pulsars (see *Nature Physical Science*, **229**, 218; 1971). K. Davidson, F. Pacini and E. E. Salpeter have now published a detailed model for the thermal X-ray emission from Sco X-1 in terms of a hot gaseous region surrounding a rotating neutron star (*Astrophys. J.*, **168**, 45; 1971). Although this model was developed before the recent flurry of activity which has resulted in the discovery of periodic variations in Centaurus X-3 and Cygnus X-1, these discoveries can, on the face of it, only add to the interest in models such as that of Davidson *et al.*

With all models of Sco X-1, the fundamental problem is the generation of thermal bremsstrahlung—the source has maintained a temperature of roughly 5×10^7 K for a decade, and it has been widely accepted that the basic energy source must be gravitational, with accre-

tion of gas onto a white dwarf releasing energy as proposed by Prendergast and Burbidge in 1968. But the difficulty of explaining the detailed flaring and flickering of Sco X-1 at both optical and X-ray frequencies, together with the presence of intermittent periodic and quasi-periodic structure in the light curves from the source, makes it desirable that other models should also be studied. Davidson, Pacini and Salpeter have looked at models in which the energy source is a rotating neutron star, the inspiration for this stemming from the apparent success of the rotating neutron star model in explaining the energy generation mechanism operating in pulsars.

The model which is most successful has a dense shell of gas surrounding the central pulsar. Supported by particular and electromagnetic radiation far beyond the speed of light circle, at a distance of roughly 10^9 cm, this cocoon is heated to the observed temperature by the radiation it absorbs. For Sco X-1, the observations require that such a shell must be quite thick, the density in the line of sight being several g cm^{-2} so that the pulsar frequency fields do not penetrate the cocoon. Plausible values for the efficiency of the heating mechanism and the mass of the underlying star suggest a pulsar period of 3 ms and a gas cloud radius of 1.5×10^9 cm. But the surface field required is only 5×10^9 G,

Infrared Variations in Seyfert Galaxies?

ALTHOUGH measurements of the infrared emission of the Seyfert galaxy NGC4151 at $11 \mu\text{m}$ do not yield any evidence of significant flux changes with time, the possibility that some flux changes have occurred cannot be ruled out. That is the principal conclusion reached by Stein and Gillett in next Monday's issue of *Nature Physical Science*.

These results are of particular interest because of their importance for the radiation mechanisms at infrared wavelengths greater than about $10 \mu\text{m}$. One model of Seyfert galaxies, for example, proposes that there is a central luminous source of short wavelength radiation surrounded by a cloud of particles that radiate at infrared wavelengths. This model would be untenable if there were fluctuations on a time scale of less than about a year. Variations of the order of a day would also place serious constraints on models involving non-thermal infrared sources.

Quite a large amount of evidence has been accumulated for time variations of the visible and radiofrequency flux from the nuclei of Seyfert galaxies and possible variations have also been reported at the near infrared wave-

lengths of about 2 to $3 \mu\text{m}$; Stein and Gillett have now extended these measurements to $11 \mu\text{m}$ with photometric measurements of NGC4151 made during the two years up to June 1971.

They have found two pieces of evidence that seem to suggest variations on a time scale of less than a month. In both cases the three standard deviation upper limit of one measurement was much lower than the lower limit of the flux a few days later. But Stein and Gillett are quick to point out that deductions from this can be misleading, and they apply several statistical checks. They find, for example, that the standard deviation of the distribution of their twenty-nine measurements is only slightly greater than the average standard deviation of the individual measurements whereas, if variations had occurred, larger differences would be expected. The numbers of measurements lying within one, two and three standard deviations of the mean of the twenty-nine measurements are also close to the numbers expected if there is no variation. Similar conclusions were reached after analysis of measurements on a comparison star (γ U Ma).

which is much lower than the surface fields required for pulsars and allows the extremely rapid rotation to persist for millions of years.

Typical time scales which Davidson *et al.* expect to be associated with the motions of the components of their model range from the pulsar period, roughly 0.003 s, to the time scale for dynamic equilibrium, roughly 3,000 s. The orbital period of the gas is 30 s, and its cooling time is roughly 0.3 s, so that it must indeed be continuously excited for the system to be stable.

Most of the objections to this model are certainly no more severe than those which apply to other models for Sco X-1. Indeed, with the exception of the rather low magnetic field, the estimated parameters agree well with observations of this and other sources, and the cocoon seems a suitable source of erratic variations of the kind detected. Two crucial tests seem relevant. One or two such objects could be detected without a drastic re-thinking of the theory, but many more sources in the Sco X-1 mould would rapidly become a surfeit. And if Sco X-1 itself does show a continuation of the decline in activity and, possibly, temperature which has been hinted at by some recent observations, then it would seem to be too short lived for the Davidson *et al.* model—this would certainly be a new development, because the downfall of most models of the object so far has been their implausibly short time scales compared with the observed lifetime of Sco X-1.

MINICELLS

Exploiting Cytochalasin

from our Cell Biology Correspondent

ALTHOUGH minicells, small enucleate *Escherichia coli*, have been used occasionally by molecular biologists, it seems safe to say that they became available too late in the game. Many questions, which might with comparative ease have been answered by experiments with minicells, had already been solved by hard work and ingenuity with nucleate *E. coli*. But animal virologists look to be more fortunate for it seems, if the experiments of Prescott, Kates and Kirkpatrick (*J. Mol. Biol.*, **59**, 505; 1971) are anything to go by, that enucleate cells obtained by exposing nucleate cells in culture to cytochalasin B can be exploited to answer many outstanding questions.

Prescott and his colleagues have used enucleate L cells produced in this way to prove quite unambiguously that the uncoating and replication of vaccinia virus does not depend on the presence of a host cell nucleus. After exposing L cells to cytochalasin B, and in some experiments to actinomycin D as well to inhibit any RNA synthesis that the cyto-

chalasin B might induce, and then washing the cultures and incubating them for 2 h in medium they added vaccinia virus and followed the fate of the infected nucleate and enucleate cells by light microscopy and autoradiography. Typical "factories" synthesizing vaccinia DNA develop in the enucleate cells and so the suggestion, originally made by Joklik in 1964, that the uncoating of vaccinia virus depends on a function coded by a host cell gene and activated by infection, has been neatly eliminated.

Joklik's suggestion stemmed, of course, from the finding that after the removal of the lipoprotein outer coat of vaccinia virus, the second step in its uncoating, the disruption of the inner "core" protein coat is specifically blocked by inhibitors of RNA and protein synthesis. This indicated that the synthesis and translation of messenger RNA was essential for the second stage of uncoating and in 1964 it seemed quite plausible that the

gene(s) in question might well be cellular rather than viral. By 1967, however, Kates and McAuslan had shown that vaccinia cores with their protein coat intact can make and release messenger and Joklik's idea was beginning to look unnecessarily complex. Prescott's results are not altogether unexpected therefore and the chief interest in what he and his colleagues have to say lies in the method they have used.

Clearly, hypotheses which implicate directly the host genome in the replication of other animal viruses might be put to the test with enucleate cells. It would, for example, be interesting to know how far the replication of RNA tumour viruses proceeds in such cells. For that might reveal whether or not the stable association between a DNA provirus, which, as Hatanaka *et al.* have just shown (*Proc. US Nat. Acad. Sci.*, **68**, 1844; 1971), is probably made in the cytoplasm, and the host genome is essential for replication of these viruses.

Transfer RNA Precursor Sequenced

In next Wednesday's *Nature New Biology*, Altman and Smith describe their latest investigations of the tyrosine transfer RNA precursor molecule of *Escherichia coli* and its maturation into tyrosine transfer RNA; their work is a remarkable example of the pitch of sophistication which has been reached in analyses of bacteria.

As Altman recently reported in *Nature New Biology* (**229**, 19; 1971) transfer RNAs, like ribosomal and probably messenger RNAs, are transcribed as part of a larger precursor molecule which is matured by cleavage. What Altman and Smith have now done is analyse the base sequence of the precursor of a tyrosine transfer RNA, prove that mutations which alter the amount of the mature RNA produced are located in a region of the precursor close to the cleavage point and detect a nucleolytic activity in extracts of *E. coli* which *in vitro* matures the precursor.

Obtaining strains which accumulate precursor involved a series of elegant genetic manipulations; with these achieved Altman and Smith obtained enough precursor to analyse its base sequence and they found several striking features. First, bases which, in mature transfer RNA, are modified are, in the precursor, unmodified. Second, it seems that the CCA sequence at the 3' terminal end of the mature molecule is present in the precursor, contrary perhaps to expectation. Third, many of the oligonucleotides obtained after partial digestion of the precursor with T₁ ribonuclease correspond to those obtained after partial digestion of the mature molecule. This situation suggests that even in the precursor the

secondary structure of the mature tRNA sequence has a clover leaf secondary structure. Finally, the 5' terminal residue of the precursor is pppGp, which suggests that transcription *in vivo* as *in vitro* is initiated with a 5' triphosphate purine nucleotide.

Altman and Smith then sequenced the precursor specified by a mutated tyrosine transfer RNA gene, which *in vivo* matures unusually rapidly. To do this, of course, they had to introduce other mutations which caused this precursor to accumulate. The sequence revealed that the mutation which leads to a rapidly maturing precursor changes a base four nucleotides removed from the cleavage point. Moreover, other mutations which cause the production of reduced amounts of mature transfer RNA but no accumulation of precursor also change bases close to the cleavage point. For these reasons Altman and Smith suggest that both the base sequence and the secondary structure at the cleavage point probably affect the rate and the precise point of cleavage and they propose models of the structure of the precursor which might account for these factors.

Finally, they have succeeded in detecting in extracts of *E. coli* a nucleolytic activity which *in vitro* cleaves the precursor at the correct point and yields a transfer RNA fragment with the same 5' terminus as transfer RNA matured *in vivo*. This finding proves that modification of certain bases in the transfer RNA segment of the precursor is not essential for proper maturation. Indeed, modification may well occur *in vivo* after the precursor has been cleaved.

HAEMOGLOBIN

Monomers Monitored

from our Molecular Biology Correspondent

THE nature and strength of subunit interactions in proteins has an intimate bearing on the interconversion between states of high and low ligand affinity, which lies at the heart of control of activity. Here, as so often, haemoglobin is the archetype. That the tetramer is in dissociation equilibrium with dimer was established in the very dawn of protein history by the old masters, Svedberg and Pedersen. More recently debate has centred on whether, at sufficiently low concentration, the dimer will in turn give rise to monomer. The general view was that at the lowest observable protein concentrations the monomer remains below the level of detectability. Schachman and Edelstein, on the other hand, reported substantial dissociation into monomers, on the basis of sedimentation equilibrium measurements with scanner optics. This conclusion has now been repudiated by Kellett and Schachman (*J. Mol. Biol.*, **59**, 387; 1971).

Scanner optics for this system offer the advantage that by operating at the wavelength of the very intense Soret absorption band of the haemoglobin, measurements can be made at concentrations as low as about 10^{-8} M in subunits. Unfortunately, oxyhaemoglobin in such dilute solution is prone to oxidize to the ferric form, which then denatures, with irreversible formation of monomeric chains. This seems to have been the cause of the previous error. Kellett and Schachman's re-evaluation was made possible by the remarkable observation that methaemoglobin formation is inhibited by the presence of EDTA, which one can only suppose to function by sequestering traces of heavy metals that would otherwise catalyse the oxidation reaction. At all events, under these conditions the equilibrium distribution in the ultracentrifuge reveals that only oxyhaemoglobin dimer is present, and the same is true of other stable derivatives, such as carboxyhaemoglobin and methaemoglobin cyanide and azide. It may be noted that Bucci (*ibid.*, **55**, 401; 1971) also recently concluded that at concentrations down to 10^{-8} M, oxyhaemoglobin gave rise to no significant amount of monomer, for a complement fixation technique, sensitive at such low levels, revealed the absence of any hybridization between two different molecular species, such as would occur were there a finite degree of dissociation beyond the symmetrical dimer ($\alpha\beta$) stage.

In an accompanying article, Kellett (*ibid.*, **59**, 401; 1971) tackles the much more demanding problem of dissocia-

tion of oxy and deoxyhaemoglobin under the influence of high salt concentrations. The difficulties are both experimental, in regard to the anaerobic manipulation of the deoxygenated solutions, and theoretical, in terms of the thermodynamics of the three-component system of protein, water and salt. Kellett has devised an analytical procedure which yields the dissociation constant for the dimer-tetramer equilibrium in terms of a buoyant weight, extracted directly from the sedimentation distribution. The upshot is that whereas the dissociation of oxyhaemoglobin to dimers (though not monomers) is strongly promoted by salt, deoxyhaemoglobin remains tetrameric even at the lowest observable concentrations in 2 M sodium chloride or iodide. Qualitatively the finding of such resistance of the unliganded protein to dissociation is not new, but Kellett's results make it clear that its dissociation constant is not less than three orders of magnitude lower in 2 M sodium chloride than that of oxyhaemoglobin. Evidently then, resistance to dissociation is a feature of the "constrained" (low-affinity) structural state. This has a number of implications in relation to functional schemes that have been propounded, and also explains the apparently anomalous failure of the cooperativity (Hill con-

stant) to diminish in high salt, when, it was thought, the protein was dissociated to dimers.

It is interesting to find that some putative evidence for haemoglobin monomers in a quite different context has now been annihilated. The haemoglobin-binding protein of serum, haptoglobin, attaches to haemoglobin at sites on the α -chains. Nagel and Gibson found that when haptoglobin was partly saturated with haemoglobin α -chains, β -chains were taken up, and this process occurred at a faster rate than when whole haemoglobin was added in place of β -chains. This result was put down to a slow dissociation of the tetramer, by way of the known $\alpha\beta$ dimers, to single β -chains, which then rapidly associated with the haptoglobin-bound α -chains. Boyd, Smith and Andrews (*Canad. J. Biochem.*, **49**, 891; 1971) prepared the α -chain-haptoglobin complex, and added to it an excess of mouse haemoglobin. By gel filtration they then separated out of the mixture, first, a haemoglobin-haptoglobin stoichiometric complex, second, mouse haemoglobin, and, finally, human α -chains. Evidently then, the initially bound α -chains are displaced. Now it is known that mouse and human haemoglobins hybridize satisfactorily. Boyd *et al.* conclude therefore that monomers do

Supermassive Rotors and Active Galactic Nuclei

A LONG standing model for quasars and active galaxies postulates the presence of supermassive differentially rotating stars, supported by radiation pressure, at the centres of these objects. In spite of other suggestions, including the possibility of pulsar-like emission from a supermassive rotating object, this relatively simple model remains a likely explanation for the energy source in quasars and galaxies. But the model is only superficially simple—although it is clear from the most naive calculations that a supermassive rotating disk satisfies the broad requirements of the primary energy supply in highly energetic sources, quantitative calculations concerning the evolution of such disks remain inadequate. It is only within the past two years that it has been shown that disks of this kind are unstable against fragmentation, and although this complicates further calculations the prospect raises some interesting conjectures regarding the ultimate fate of the disks. E. E. Salpeter, of Cornell University, has now provided an outline of the successive fragmentations which should accompany the slow evolution of a supermassive star. This approach, presented in next Monday's *Nature Physical Science*, raises

several questions which point the way for further development of the model.

It seems that, so long as the disk is radiation dominated, fragmentation must occur. This process will be repeated in the subdisks formed, until small objects supported by hydrodynamic forces are formed. Because the descendants of the original disk will form a hierarchical grouping after several fragmentations, radiation escaping from the system will suffer successive redshifts as it leaves first the subgroup from which it originates, then the group of subgroups in which that subgroup lies, and so on.

One source of electromagnetic radiation could be pulsar-like activity, which would occur just before the fragmentation of any disk which carries a magnetic field.

The major difficulties in extrapolating this simple picture to more realistic models concern differential rotation and the detailed nature of the gravitational and pulsar radiation emitted. Gravitational radiation, for example, might result in selective changes in the orbits of the fragments until a stable, symmetric pattern with no gravitational radiation is formed.

not enter into the scheme at all. Whereas β -chains can add directly to the complex; the reactive species in haemoglobin is the $\alpha\beta$ dimer, and the slow step is then the dissociation of each α -chain from the haptoglobin to allow an $\alpha\beta$ unit to bind in its stead.

UPPER MANTLE

Tasman Conductivity

from our Geomagnetism Correspondent

THE Tasman Geosyncline, which lies along the southern part of the east coast of Australia, is a long structure generally some 200–300 km wide and includes two major topographic features, the Australian Alps and the Main Dividing Range, as well as the three major cities of Canberra, Melbourne and Sydney. Geologically it comprises folded and metamorphosed rock belts of Ordovician to mid-Devonian age and includes the major granite batholith which forms Mount Kosciusko, at 2,219 m the highest point of the Australian continent. Seismic refraction studies have shown that the crustal rocks beneath Canberra are about 40 km thick but thin to about half of that in the coastal region. The eastern boundary of the geosyncline is not particularly clear although seismicity extends to the continental shelf. The western boundary, on the other hand, is rather different; seismological studies have indicated quite a clear boundary to the current tectonically active belt. Little, however, is known about the region immediately to the west of the geosyncline in spite of its proximity to the three principal Australian cities. Outcrops are poor there and the only major structure mapped, and that superficially, is a sedimentary basin.

To learn something about the behaviour of the upper mantle, both beneath the geosyncline and to the west of it, Tammemagi and Lilley (*Geophys. J.*, **22**, 505; 1971) have made four magnetotelluric soundings at 150 km intervals roughly along the latitude of Canberra—at Moruya on the east coast, Spring Valley near Canberra almost on the crest of the belt, Wagga Wagga slightly to the west of the geosyncline and Griffith well to the west. The magnetotelluric method itself is not new, although in this case automatic equipment was operated at each site for at least four weeks, a rather longer than usual period, to ensure that enough events were recorded. In spite of the long history of the method only a handful of magnetotelluric measurements have ever been made in Australia and none at all in what could be regarded as Australia's best known geological region.

Three events or periods of activity, including at least one magnetic storm, were frequency analysed for each site and for signals in the period range 10^3 to 10^4 ; and the usual curves of apparent resistivity (of the upper mantle) were constructed. As is well known, the interpretation of magnetotelluric data is frequently hampered by the inherent ambiguities, especially where the frequency range is limited and where the curves of apparent resistivity show anisotropy. In this particular work, three of the four stations were troubled by anisotropy. At Moruya, on the east coast, ocean-induced anisotropy is clearly present. The data are most consistent with a two-layer conductivity model although because of the anisotropy a large number of detailed models within this classification are possible. At Wagga Wagga and Griffith the situations are similar except that three-layer models are more appropriate. But at Spring Valley, close to the crest of the geosyncline, there is no anisotropy. A two-layer conductivity model is again most appropriate here although

a range of electrical conductivity structures is possible.

But although anisotropy and the limited frequency range have prevented Tammemagi and Lilley from making detailed conductivity models, one important fact has emerged from the data—this is that in the depth range 100–300 km the electrical conductivity under the stations Moruya and Spring Valley is an order of magnitude higher than that under Wagga Wagga and Griffith.

In other words, the conductivity is higher under the Tasman Geosyncline than under the region to the west of it. In view of the correlation between electrical conductivity and heat flow discovered in Utah and Colorado by Reitzel *et al.* (*Geophys. J.*, **19**, 213; 1970), Tammemagi and Lilley interpret this as a temperature difference in the upper mantle between the two regions. Using the relationships presented by Tozer (*J. Geomagn. Geoelect.*, **22**, 35; 1970), they calculate that at a depth of about 200 km the temperature difference should be about 200° C; but whatever the precise figures, the existence of a

Helping Human Adenovirus to Replicate

SEVERAL human adenoviruses infect only abortively African green monkey kidney cells—the virus particles penetrate these cells and are uncoated; viral RNA is made, the viral genome replicates and at least some viral proteins are synthesized, but very few progeny virus particles are produced. In these monkey cells it seems that either some crucial viral proteins are not synthesized or the transport of the viral proteins from the cytoplasm where they are made to the nucleus where progeny virus particles are assembled is blocked. By coinfecting the cells with simian virus 40, however, these blocks are removed and the yield of adenovirus is greatly increased. In other words, SV40 provides some helper function which permits human adenoviruses to complete their replication in non-permissive monkey cells. In next Wednesday's *Nature New Biology*, Henry, Slifkin and Merkow report experiments which shed some light on this intriguing phenomenon.

Henry *et al.* infected with human adenovirus 2 coverslip cultures of non-permissive African green monkey kidney cells and permissive VERO and HeLa cells, which support the complete replication of this virus. Some of the monkey cell cultures were also coinfecting with SV40 either immediately or within 24 h and 5-fluorodeoxyuridine (FUDR) was added to some of the coinfecting cultures. This drug blocks the synthesis of "late" SV40 proteins but does not impair the SV40 helper func-

tion(s). At 24 or 45 h after infection 2 min pulses of labelled amino-acid were fed to the cultures which were then chased with cold medium. The synthesis and transport of protein made during the pulse could then be measured by autoradiography.

In brief Henry *et al.* found, using an immunological precipitation assay to differentiate between viral and host proteins, that some 60–70 per cent of the viral protein made in permissive cells is transported into the cell nuclei. In the monkey cells infected by adenovirus 2 alone only about 20 to 30 per cent of the protein labelled during the pulse is in the nucleus some 2 h later, the remainder is still in the cytoplasm concentrated about the periphery of the nucleus. By contrast, in monkey cells coinfecting with SV40 and exposed to FUDR about 50 per cent of the adenovirus protein is transported to the nucleus. Further, the amount of viral protein made in the coinfecting cells is significantly greater than the amount made by the cells infected with adenovirus alone.

It seems therefore that some SV40 gene product(s) not only enhances the synthesis of adenovirus proteins but also promotes the transport of these proteins into the nucleus where they are assembled into capsomeres and capsids. It is probably this second helper function which is chiefly responsible for increasing the yield of progeny virus particles; its molecular basis, needless to say, awaits elucidation.

temperature contrast is presumably reflected by, either as cause or effect, the tectonic activity within the geosynclinal region.

The possible presence of temperature variations naturally raises questions about the heat flow in the area although, unfortunately, very few measurements have been made. There is, however, a published heat flow value of $2.0\text{--}2.2 \mu\text{cal cm}^{-2} \text{ s}^{-1}$ at Spring Valley and one of $1.1\text{--}1.2 \mu\text{cal cm}^{-2} \text{ s}^{-1}$ at Moruya which suggest that towards the middle of the geosyncline at least the heat flow is higher than average in keeping with the enhanced mantle temperature. By the same token, heat flow to the west should be lower, though not necessarily lower than the world average. The only heat flow determinations to the west are several hundred kilometres away at Broken Hill which is itself in a tectonic region, the Adelaide Geosyncline. Tammemagi and Lilley thus predict a heat flow trough between the two geosynclines and, should such a trough be proved, look forward to confirmation of their interpretation of the magnetotelluric data.

CENTRAL NERVOUS SYSTEM

Useful Brain Cells

from our Cell Biology Correspondent

BIOCHEMISTS debating whether or not to take the plunge and tackle some of the biochemical problems posed by the development and differentiation of the central nervous system should be encouraged by what Seeds has to say in the latest issue of the *Proceedings of the US National Academy of Sciences* (68, 1858; 1971). For he has shown that dissociated mouse embryo brain cells, when put in rotation cultures, form themselves into aggregates which continue to undergo at least some of the biochemical changes that occur during the normal postnatal development of the intact mouse brain.

Seeds chose to follow the activity in his aggregates of three enzymes, choline acetyltransferase, acetylcholinesterase and glutamate decarboxylase, which are associated with the process of nerve transmission and which greatly increase in activity in the mouse brain during the first three to four weeks of postnatal life: the central nervous system of mice is unusually poorly developed at birth and both the formation of synapses and the growth of dendrites as well as biochemical development occur after birth. Seeds finds that the time course and pattern of increase in the activity of choline acetyltransferase and acetylcholinesterase in intact brains and in the aggregates of embryonic brain cells in culture are encouragingly similar.

Moreover, morphological differentiation also occurs in the aggregates; Seeds mentions, for example, the presence of numerous synapses. Although nobody suggests that these aggregates can be used to solve all the questions posed by the development of the central nervous system they at least offer a most promising starting point.

Brain cells also have attractions for biochemists interested in cytoplasmic microtubules, and Bryan and Wilson (*ibid.*, 1762) report that these organelles in embryonic chick brain cells are polymers of two species of protein not one. Using colchicine as a label, several groups have isolated microtubule subunits from a variety of different cells. The best characterized of these, isolated from adult human brain, appears to be a dimeric molecule comprising two polypeptide chains weighing about 55,000 to 60,000 daltons, and it has generally been assumed that the two polypeptides in each dimer are identical. Bryan and Wilson, however, find that the two polypeptides of the dimer protein of chick brain microtubules can be separated by polyacrylamide gel electrophoresis after they have been reduced and acetylated. Furthermore, although these two polypeptides have the same

molecular weight, about 55,000, they differ significantly in their content of six amino-acids.

From this and similar studies of microtubules from other sources Bryan and Wilson conclude that the basic subunit of these organelles is an $\alpha\beta$ dimer, rather than an $\alpha\alpha$ dimer or a mixture of $\alpha\alpha$ and $\beta\beta$ dimers.

CHEMOSYSTEMATICS

New Committee

A JOINT Committee on Chemotaxonomy has recently been formed by the International Union of Pure and Applied Chemistry (IUPAC) and the International Association for Plant Taxonomy (IAPT) to look into all aspects of the organization of international collaboration in chemosystematics. The committee invites comments from interested persons; these should be sent to the chairman (Dr W. F. Grant, Genetics Laboratory, MacDonald Campus of McGill University, Ste. Anne de Bellevue 800, Quebec, Canada) or to the secretary (Dr T. Swain, Royal Botanic Gardens, Kew, Richmond, Surrey, England).

Polyadenine and Haemoglobin Messenger

LAST year, Lim and Canellakis (*Nature*, 227, 710; 1970) discovered tracts of polyadenine, some fifty to seventy bases long, in the ribonuclease digests of purified rabbit globin messenger RNA. Since then similar tracts of poly A have been detected in RNA fractions, isolated from the polyosomes of HeLa and ascites cells, which are also presumably messengers. It seems therefore that the presence of sequences rich in adenine may be a general feature of many, perhaps even all, eukaryotic messengers. Because of their restricted sequences it has been generally assumed that these tracts are not translated but occur on either, or on one side of the coding sequence in the messenger, and in *Nature New Biology* next Wednesday Burr and Lingrel report direct evidence that this is the case.

After labelling the 3' terminal residues of rabbit globin messenger RNAs with tritiated borohydride they digested the molecules with either T_1 or pancreatic ribonucleases and isolated the 3' terminal fragments. After T_1 digestion these fragments elute from 'Sephadex' columns in the region at which oligonucleotides with eight phosphate residues appear. The terminal

base of this fragment is adenine. The 3' terminal fragments obtained after pancreatic ribonuclease digestion, which cannot contain pyrimidines, are five and six bases long and, because the first guanine residue cannot be within nine bases of this terminus, Burr and Lingrel conclude that these messengers have at their 3' termini a tract of five or six adenine residues and from the known amino-acid sequences of the rabbit globins it is clear that these bases are not translated.

The sequences Burr and Lingrel have analysed are considerably shorter than those detected by Lim and Canellakis, and there are several possible explanations of this difference. Much longer poly A sequences may exist at the 5' ends of these molecules; the short sequence Burr and Lingrel have determined may be part of a much longer adenine rich tract or the poly A tracts may be labile and partially lost either during the extraction of the messenger or indeed *in vivo* as the molecules age. Needless to say, the function of poly A in eukaryotic messengers is as yet a matter of speculation and, as anybody who reads what Burr and Lingrel have to say will realize, there is no shortage of ideas.

The Doomsday Syndrome

JOHN MADDOX

THE past few years have been quite remarkable for the way in which there has flourished a rich vein of commentary on the place that science and technology occupy and should occupy in the modern world and for the way in which the underlying theme of this commentary has tended to be that science and technology will quickly bring society to its knees. Even quite short memories are sufficient to show that this development flies in the face of the view of the social function of science and technology that was current and commonly accepted as recently as twenty years ago. Then, and for the best part of a century before, it was supposed that social progress must rest on technical progress of some kind or another. The great philanthropists of the nineteenth century, for example, were constantly linking their schemes for liberating people from unpalatable and degrading occupations with the hope that mechanical alternatives would be found for jobs of work that were manifestly unsuitable for people. They were buoyed up by the encouragement of the nineteenth century liberals and by the legacy of the earlier utopians—people like Godwin, for example—who took the view that there would be no limit to the achievements of the human spirit so long as people could be liberated from the need for drudgery so as to keep alive and from the restraints of the environment in which they found themselves and which was, almost by definition, hostile. It is hard to imagine a more complete contrast than that between the way in which technical developments used to engender optimism and the way in which the prospect of technical innovation has now become a kind of threat.

The signs of this new attitude are now plain to see. First of all, the commentators are conspicuous and what they have to say is arresting. Several arguments have been used in the past few months to heighten the sense of impending disaster. The rapid growth of the world's population is, for example, one of the most common. And, of course, it is by now well known that the population of the world cannot continue to grow at the present pace for decade after decade. On present estimates, the world's population probably falls only little short of 4,000 million, and so many of the people now alive are either children or still in the child-bearing age group that it seems unlikely that contraceptive devices, however fully and ardently employed, will allow the world's population to stabilize at less than 7,000 million. This population should be reached by the turn of the century. It goes without saying that the most spectacular growth will be in the countries still at an early stage of their economic development so that, by the year 2000, countries such as Brazil and Indonesia will be roughly comparable in the size of their populations with the United States and the Soviet Union.

The prospect of this pattern of growth has apparently caused great consternation in the past few years. The most vociferous prophet has been Dr Paul Ehrlich, who has held quite openly that what he calls "the population bomb" must spell disaster for society at large, not merely for those unfortunate communities where the pace of growth is greatest. And what kind of disaster? Dr Ehrlich says quite categorically in his best-known book that the world will run out of food—that the mid-seventies will be times when famine is rife and when tens of millions of people will die each year of starvation. This view, itself an echo of the doctrines which Malthus put forward in the earlier versions of his distinguished tract

on population, has been taken up by all kinds of commentators, from people like Lord Caradon to a host of popularizers. In the circumstances, it seems almost an indecent question to ask whether disaster is as near as all this talk suggests.

In the first place, however, it is a surprising fact that the population of the world has grown less quickly than its production of food. The Food and Agriculture Organization of the United Nations, in its time one of the most fervent beaters of the calamity drum, has kept statistics of food and population going back to the early nineteen thirties. The remarkable feature of these statistics is the way in which food and population appear to march hand in hand. Although there are obviously ups and downs in the statistics of food production, chiefly because of fluctuations of climate from one year to another, the production of food per head of population has actually increased by five per cent between the late thirties and the late sixties. And now, as everybody knows, many developing countries are faced with the prospect not of famine but of the social and economic problems of marketing harvests of cereals which so far exceed the past performances of peasant farmers that it is necessary quickly to develop the commercial machinery long familiar in the agriculture of nations in which subsistence agriculture has long since disappeared. In other words, whatever the prophets of doom may say, it does seem that famine will not be the most obvious cause of calamity in the next decade. On this point, it appears, the doomsday view seems entirely inappropriate.

To say all this does not imply, of course, that the pace of growth of the world's population brings no unfortunate or unforeseen consequences, or that nothing needs to be done about it. It is, however, important that these problems should be seen in their historical perspective. Students of population are impressed with the way in which it seems to have taken the human race from the beginning of its evolution two million years ago to the beginning of the nineteenth century to reach 1,000 million, a further century to add a second thousand million to the total roster and merely half a century since then to double the population of the world. On the face of things, this is much more than exponential growth—the time constant is continually decreasing. Yet there are good reasons for thinking that the growth of a human population is much more complicated and subtle than the growth of a population of bacteria in a set of Petri dishes.

In Britain, for example, the mediaeval population was kept in check by a succession of infectious diseases, from plague to influenza. The agricultural revolution in the eighteenth century so increased the productivity of the land that the population was able to grow quite quickly, and there is plenty of evidence that from the second half of the eighteenth century on, the birth rate in Britain was much less than it might

The pages which follow are intended as contributions to the wider understanding of the interest and excitement in contemporary science on the occasion of the 133rd annual meeting of the British Association at Swansea from September 1 to 8.

have been if all women of child-bearing age had produced their full biological quota of children. In other words, it looks as if it is quite possible for a growing population to limit its rate of growth even without the full apparatus of modern contraceptive techniques. The way in which the population of Britain took another spurt forward in the second half of the nineteenth century is a little more puzzling. Part of the explanation is that children died less frequently in infancy, and therefore survived to have children of their own, but it also seems that comparatively prosperous families were able deliberately to decide to have larger families.

This brief history is relevant to what is happening now elsewhere in the world. First of all, the rapid growth of the world's population in the past thirty years is almost directly attributable to the way in which better medicine and better insecticides (used against the mosquitoes which carry malaria, for example) have enabled hard-pressed communities to keep their children alive until middle age. In some quarters, quite misguided, this is held to be a proof that medicine and insecticide technology can be "blamed" for the growth of the world's population. The truth is, of course, quite different. In present circumstances, there is good reason to think that one of the conditions that must be satisfied before rapidly growing populations take the steps to limit the pace of growth which seem to have taken place in the eighteenth century in Britain is that there must be a still further improvement of medical care in infancy. In South America, India and South-East Asia, it seems to be quite commonplace that mothers when asked how many children they would like to have provide higher answers when the infant mortality is high. For reasons which are obvious and which cannot be dismissed, they take the view that they must have enough babies to make sure that one at least will grow up to be an adult male. But it is clear that prosperity is another condition to be met, as recent experience of a falling birth rate in places such as Hong Kong and Singapore bears out. As it happens, all this is quite familiar and, to say the least of it, it is surprising that the prophets of doom have made so little of this to them unpalatable evidence.

What this implies is that there is every reason to expect that social forces, already apparent in the recent demographic history of Britain, will in due course help to control the populations of developing countries. To be sure, it is important that steps should be taken to encourage these developments. If the true objective of those who preach calamity is to bring about a stable solution, they would be much better advised to advocate those kinds of social innovations that might encourage stability than to complain that famine is only just around the corner. In doing so, of course, they would be driven to recognize that one of the ironies of the contemporary problem of population growth is that the countries where the population is growing most quickly are so compelled to invest resources in schools and medical facilities that they have very little for the kind of industrial investment that might in the long run make them economically prosperous, itself a precondition for population stability. The moral in this is no doubt that foreign aid of the right kind could make an important contribution to the stabilization of the world's population. The modern equivalent of Malthus's position is that population growth brings impoverishment. It is irresponsible of the doomsday men to shirk these subtleties, and to settle for a kind of biblical choice between survival and calamity.

The alarm which attends the public discussion of the population problem crops up everywhere. Pollution, for example, is now so widely held to be a threat to the survival of orderly society that there is a danger of a serious distortion of social priorities. DDT and the other chlorinated hydrocarbons which have functioned as insecticides since the war are good examples. The materials are persistent, and are toxic to animals which concentrate them. It seems to be well established that birds such as peregrine falcons and eagles are par-

ticularly affected by these insecticides. At the same time, however, nobody would suppose that the damage done to eagles in, say, North America by insecticides is anything like as great as the damage caused by the illegal shooting of eagles by farmers in the Dakotas and elsewhere in the Middle West of the United States. And, moreover, there is no evidence at all that these insecticides have been damaging to people. To be sure, experiments in the United States, admittedly inadequate and only preliminary in character, have shown that the chlorinated hydrocarbons in very large doses can increase the incidence of tumours in mice, but there is as yet no sense of what bearing this may have on human health, although it does seem clear that the long range effects of DDT and similar materials are unlikely to be more damaging than a great many other familiar chemicals in the environment, and certainly less damaging than some of those which have in the past few decades been eliminated altogether—the use of carbon disulphide in vulcanizing plants, for example. Although the persistence of DDT argues for close control of the way in which it is used, it is surely entirely misguided that an over-zealous scorn for DDT and related chemicals should rob the world at large of the benefits they may have in the prevention of malaria, in the improvement of precisely the kinds of agricultural productivity that bear on the problem of how to feed the growing population and which are, in any case, a good deal less damaging to the environment than some of the materials—sodium arsenate, for example—which they replace. To be sure, it is always possible to hope that more subtle methods of controlling insect populations, ranging from the deliberate infection of insect populations with parasites to the use of pheromones for baiting insect traps, may in due course help to control populations of unwanted insect pests without causing serious side effects. But it is only proper that those who advocate the abolition of DDT on the grounds that it will bring calamity should warn the agriculturists whom they would deprive that biological control is still a technique applicable only in particular circumstances.

Population and pollution are, however, only two of the spectres which perplex commentators on the relationship of science and technology with society. Other people fear that carbon dioxide in the atmosphere will eventually produce such a powerful greenhouse effect that the Antarctic ice cap will be melted but, in reality, nobody can be sure that the effect will be that predicted and in any case the accumulation of carbon dioxide in the atmosphere is by no means the inexorable process that the doomsday men suggest. Other recent fears—the current alarm about the way in which exhausts from supersonic aircraft may destroy the ozone layer in the atmosphere and thus expose people to dangerous doses of ultraviolet light—are similarly unfounded, but this does not undermine their usefulness in the hands of those who seem anxious to demonstrate that the world is about to end.

In the long run, the most interesting feature of the present wave of gloomy speculation is likely to be the sudden emergence of what will, no doubt, be considered to have been a wave of fashion. Most probably, the real threat of global catastrophe posed by nuclear weapons in the forties and more especially the fifties is a part of the explanation why there should now be such concern: that global calamity might arrive. But, paradoxically, the threat of nuclear war on a catastrophic scale is now in retreat. That the fifties should have left behind such a sense of bewilderment about the way to strike a balance between the useful applications of technology and the unwanted side effects which it can produce is a social phenomenon of great interest. Is it possible that the explanation is to be found in the sharp contrast between the doomsday view now prevalent in advanced societies and the optimism about science and technology still to be found where prosperity is, for the time being, less in evidence?

An extract from an address to Section X of the British Association.

Science, Statistics and Society

A. W. F. EDWARDS

Department of Medicine, University of Cambridge

Statistical techniques primarily developed for the natural sciences are finding increasing application in the social sciences. At a time when there is little agreement among statisticians about the proper logical basis of statistical inference, the use of equivocal methods for the interpretation of scientific data of social relevance, and the misunderstanding of some of the basic tenets of inductive inference, may have grave social consequences.

STATISTICAL theory has been developed, principally over the past hundred years, in response to the needs of astronomers and physicists and, more recently, to the needs of medical, biological, and agricultural research workers. The body of methods which has thus been created is now increasingly finding application in the social sciences, with the result that techniques originally invented for the reduction of astronomical observations, or for the testing of hypotheses in controlled experimental conditions, are being applied in areas which are socially sensitive. But are they the appropriate techniques? Are they sufficiently clearly understood to contribute positively to the analysis of data in the social sciences, or do they obscure the train of thought by substituting rhyme for reason?

As the "centre of gravity" of research moves from the physical, through the biological, to the social sciences, statistics (in its inductive, as opposed to purely descriptive, role) becomes of increasing importance as the filter through which the information contained in the data will pass. But in the minds of most non-statisticians this filter is a black box, filled with mysteries, even though provided with labelled inputs and outputs. The black box has recently adopted the dimensions of a computer centre, complete with packages of statistical programs available on demand to all who conform to their input requirements. Though admitted to be full of mysteries, it is widely supposed that the black box operates on principles generally accepted by statisticians. It comes as a shock to discover that the statistical world is in a ferment of disagreement over these principles. Not even does the concept of a null hypothesis and the associated simple test of significance, that hallmark of statistical numeracy demanded by the editors of progressive journals, pass unchallenged.

Controversy in Statistical Inference

As an example, consider the following question set in a university examination for a BA in Sociology¹:

"Let μ be the true IQ of student X. It is known that his scores on test Y are normally distributed with mean μ and standard deviation 4.00. The test is administered to him on occasion Z and he scores 125.

- (i) What is the maximum-likelihood estimator of μ ?
- (ii) Test the hypothesis that $\mu = 120$.
- (iii) Comment upon the statistical bases of the methods used to solve (i) and (ii)."

A well informed candidate might answer as follows:

"(i) The maximum-likelihood estimator of μ is the sample mean; in this case the estimate is therefore 125.

(ii) See (iii).

(iii) The statistical basis of (i) is in dispute. Some hold that it is reasonable to contemplate a 'true' value, and that the maximum-likelihood estimator is to be preferred because it has optimum properties under repeated sampling, too complex to be discussed here. Others hold that the concept of a 'true' value is vacuous, and that the maximum-likelihood estimate is to be taken because it offers the best fit to the data, in the sense of maximizing the likelihood. It should be noted that still others reject the use of maximum-likelihood on either justification.

"There are at least six points of view on (ii):

- (a) The question does not admit a statistical solution.
- (b) A test may be constructed on Neyman-Pearson principles, according to which the hypothesis will be accepted or rejected, as in a decision procedure, the justification being in terms of repeated sampling.
- (c) A test may be constructed on Fisher's principle in which the probability of the observed result, or one more extreme, on the basis of the null hypothesis, is used to judge whether that hypothesis is to be rejected. It might be 'one-tailed' or 'two-tailed'.
- (d) The fiducial distribution of μ may be found, and a statement of fiducial probability made about the null hypothesis.
- (e) A Bayesian position may be adopted, in which a prior probability distribution for μ , representing knowledge about μ prior to the observation, is compounded with the likelihood function derived from the observation, to form a posterior distribution for μ . There is no complete agreement over a suitable prior distribution, nor over the most suitable form of test once the posterior distribution has been found.
- (f) A pure likelihood position may be adopted, in which the

null hypothesis is compared with the most likely alternative hypothesis using the likelihood ratio.

"In view of the unresolved arguments over the applicability of these methods, I am unable to answer (ii) to my satisfaction, though it should be noted that since the distribution is specified as normal, some of the methods would give very similar answers, a circumstance that has helped to obscure important logical distinctions."

There is, of course, no time to teach such a critical approach. "In the behavioural sciences in general the overwhelming practice both by researchers and those responsible for statistical training has been to ignore the issues raised by the critics and to continue doing things as before."²

In an earlier article³ I examined some of the logical problems of the conventional approaches, concluding in favour of the use of likelihood as the basis of the calculus for inductive inference⁴. In the present article I air some of the more elementary points in probability and statistics which are often misunderstood, hoping thereby to encourage those who already harbour suspicions about many conventional arguments not to allow themselves to be dazzled by statistical brilliance.

The Null Hypothesis

No argument should be accepted simply because it relies on some orthodox statistical practice. Statistical techniques need to be viewed as aids to, and not substitutes for, thought. They will not help if they are not clearly understood, and they may seriously hinder. Take, for example, the null hypothesis, widely acclaimed in the social sciences as the *sine qua non* of objective data analysis, but viewed with suspicion by others as providing "the basis for the simplest decisional procedures which have so effectively colonized the more gullible sciences"⁵. Not so many years ago responsible people arrived at sensible conclusions from bodies of data by an intuitive process that can only be described as "judgment", in which their training, their experience, and their unique way of thinking each played a part. Such a subjective approach is now viewed with scorn; instead it is argued that we should adopt the objective statistical approach of setting up a null hypothesis and testing it. But how are we to choose the null hypothesis?

A notorious example is provided by the current debate on the inheritance of intelligence, in which it appears at first sight that a suitable null or "empty" hypothesis is that intelligence in man contains no inherited component. But since data on the inheritance of intelligence are difficult to acquire, and the subject of much special pleading, a "rejection of the null hypothesis at the 1% level of significance" is hardly to be expected. Repeated non-rejection of the null hypothesis is too easily interpreted as indicating its acceptance, so that on the basis of no prior information coupled with little observational data, the null hypothesis is accepted.

Far from being an exercise in scientific objectivity, such a procedure is open to the gravest misgivings. What used to be called prejudice is now called a null hypothesis, and what used to be called judgment is now called prejudice. For may we not, with equal absurdity, adopt the null hypothesis that all variations in intelligence are genetic in origin? Thus we find that not only are the statistical methods appropriate for testing a null hypothesis in dispute, but the very concept of such a privileged hypothesis is a potential danger.

If we base our actions—such as social policies—on such an approach, we lay ourselves open to the tyranny of null hypotheses, and the prejudices of those who choose them. Thus it would be "natural" to advance the null hypothesis that family allowances have no effect on family size, but—in the absence of evidence on the point—it would be absurd to act as if this hypothesis were true just because it has been designated "null". A further example is provided by discussions of whether violence and obscenity on television corrupt: the absence of positive evidence is sometimes taken

to justify the assumption that they do not.

In the more abstract world of pure scientific research, in which such statistical techniques evolved, questions of prejudice less frequently arise, and are generally less important when they do. But in other fields the adoption of methods which are so open to misuse is to place a dangerous weapon in uncritical hands: "The enthusiast whose sole interest in science is to find support for a preconceived social policy, not only deceives himself, and perhaps others who are not aware of his bias, but inevitably brings some discredit upon all other workers in the same field"⁶.

Moral and Metrical Equality

One of the characteristics of our age is a preoccupation with equality. In its moral sense, "the condition of being equal in dignity, privileges, power, etc., with others"⁷, surely no justification is needed, yet there is a tendency to seek one by predicating moral equality on "metrical" equality, "the condition of being equal in quantity, amount, value, intensity, etc."⁷. There is thus often a strong motive for assuming metrical equality solely on the basis of no evidence to the contrary, an assumption to which some statistical theories lend support.

In the approaches which depend on the null hypothesis, scientific respectability is apparently conferred on the concept of equality, for if we have no information on the variability of a statistical population, an apparently "empty" hypothesis is that all the members of that population are the same in respect of the attribute in question. On many socially important problems it is regarded in some quarters as irresponsible to challenge the null hypothesis with data, yet in truth the hypothesis itself is an exceptionally strong one in most applications, one which no thinking scientist would accept were it not for the influence of current statistical practice.

In the Bayesian approach a similar role is played by Laplace's *Principle of Indifference* applied to beliefs. Ignorance is represented by an equality of probability, a postulate which many cannot accept because they hold that a probability statement is invariably informative, and that the only permissible expression of ignorance is no statement at all⁸.

The predication of moral on metrical equality, and the unwarranted assumption of the latter, are weak points in a common argument; the first is an example of the best causes attracting the worst arguments, whilst preoccupation with the second, and hence a concern with relative rather than absolute standards, heralds an inflation which is not only monetary.

The assumption of equality may indeed have precisely the opposite effect to that which it is intended to have by its well meaning proponents. Thus, returning to the question of intelligence, if we assume that each child has the same intellectual potential whatever his parentage, and if this turns out to be a false assumption, those at either extreme of the Mendelian lottery will not receive the special treatment necessary to increase social equality. As has been observed on more than one occasion recently, any attempt to allocate university places or degree classes in such a way that each group (be the grouping by class, race, creed, sex, or any other category) is represented in proportion to its numbers may have most undesirable effects. For should any such group in fact be less well endowed genetically than another, it will soon become common knowledge that a first-class degree (say) is of a lower standard when the holder is a member of that group. Membership of a particular group will then be seen to be relevant to the assessment of the individual, which is precisely the situation we wish to avoid. Perhaps these matters were more clearly understood before the advent of statistics. In the sixteenth century it was well known in Cambridge that the intellectual reputation of Trinity College was founded on the fact that its fellowships

were not restricted by place of birth: in most other colleges there was a restriction on the number of fellows from each county, in an attempt to ensure national equality⁸. The system fell into disuse.

Causation, Correlation, Classification

The confusion of thought contained in the failure to distinguish the concepts of causation and correlation is so elementary that it would not need stressing were it not for its grave consequences. It is a distinction which statisticians are invariably at pains to make clear, and hence differs from those misconceptions which have their origin in statistical theory. The confusion of the two concepts is a major source of prejudice, especially in race relations.

At the same time as techniques, now regarded with suspicion by many statisticians, are sweeping through the social sciences, some of the more straightforward consequences of probability theory are neglected. Perhaps the most important of these is the fact that there are two ways by which we may acquire information about an individual. Suppose I speak to you on the telephone and we arrange a meeting, never having met each other before. I know from your voice that you are a man, and from your name that you are an Englishman. Even though this is all the information I possess about you, it enables me to make a statement about your height, thus: the probability that you are between 5 feet 8 inches and 5 feet 8.5 inches tall is 0.8, if I take as my datum the fact that Englishmen are normally distributed in height with a mean of 5 feet 8 inches and a standard deviation of 2.5 inches.

Such a statement is poised half way between complete ignorance and complete knowledge of your height; that is why it is phrased in terms of probability: "It has been often recognized that any probability statement, being a rigorous statement involving uncertainty, has less factual content than an assertion of certain fact would have, and at the same time has more factual content than a statement of complete ignorance. The *knowledge* required for such a statement refers to a well-defined aggregate, or population of possibilities within which the limiting frequency ratio must be exactly known. The necessary *ignorance* is specified by our inability to discriminate any of the different subaggregates having different limiting frequency ratios, such as must always exist"⁹.

The important point is that my statement refers to *you*. It is not true to assert—to take a recent example from the discussion on education—that "characteristics of social groups have no implications for the kinds of individual differences which are the concern of education"¹⁰. For such a statement to be valid, either the social groups must not differ in respect of the characteristic, or the logic of probability theory is at fault. The first point needs to be demonstrated. The process of induction outlined above is very old: indeed, the whole point of giving objects names at all is to classify them, and hence to convey information about them. If I say "I own a dog" you know certain things about the animal I keep in my house, though you have not seen it.

The second way in which we acquire information about an individual is by direct measurement. Thus, when I meet you, I may judge your height "to the nearest inch", or, if I carry a measure, "to the nearest tenth". But this new information does not annihilate the old; indeed for over two hundred years we have possessed a theorem which enables us to combine information from these two sources: Bayes Theorem.

The technical details of the theorem cannot now detain us, but reading sociological articles leaves a distinct impression that this most important theorem has not yet penetrated into the thinking of social scientists. In their concern to treat the individual as an individual they deny the relevance of group characteristics; but this is to deny Bayes Theorem, and hence to suppress some relevant information. Objectivity is not to be achieved by the suppression of information.

One of the consequences of the theorem is that, as individual measurements become more reliable, they become more important relative to the "prior" information derived from group characteristics, as is to be expected: a doctor who books a maternity bed for one of his patients on the basis of standard tables for the duration of pregnancy is making the best use of the meagre information at his disposal; but when the patient telephones to say that labour has started, the new information far outweighs the old.

The problem of the applicability of Bayes Theorem is a deep one. It is clearly appropriate in scientific situations such as those described above, where information must not be discarded. But it is excluded when considering criminal cases in courts of law, in which guilt is pronounced only if it is "beyond all reasonable doubt". If the direct evidence is so meagre that the prior information, such as provided by previous convictions for the same type of offence, is relatively important, then there will still be "reasonable doubt" about the guilt of the accused. Besides which, justice requires that every case be decided "on its merits", since the accused may be a reformed character for whom evidence of previous convictions would be irrelevant. He should have the benefit of the doubt.

It is in the assessment of the merits of the individual that the operation of Bayes Theorem needs the closest scrutiny. Thus in considering applicants for a post, we should ideally evaluate the suitability of each and every one with such diligence that any prior information is of negligible importance in our final decision. In practice, the effort is too great. We therefore invent qualifications, such as degrees and certificates, and only consider applicants who hold qualifications we consider appropriate: professors are almost invariably appointed from among graduates (George Boole was an exception). An inevitable cycle then starts. From being an indication of suitability for a post, a degree becomes a prerequisite. It thus becomes desirable in itself; demand exceeds supply; expansion occurs to meet the demand, and the supply is soon more than adequate. The surplus graduates are aggrieved that they cannot find employment of the kind that graduates formerly enjoyed. An inflation has taken place because the original purpose has been partially forgotten, and the cry is raised for the abolition of the system.

The consequences of misunderstanding some of the basic tenets of inductive inference, and of uncritically accepting and applying statistical recipes whose foundations are still in dispute, are grave. Quantitative work in the social sciences suffers and inappropriate modes of thought are introduced into otherwise rational discussion.

We need not applaud the extreme reaction of Hogben¹¹ to the penetrating advances that have been made this century towards a better understanding of statistical inference, but we would be foolish not to heed his warning: "Laboratory experiments will have to stand on their own feet without protection from a façade of irrelevant computations. Sociologists will have to use their brains. In my view, science will not suffer".

¹ Quoted by Rosenbaum, S., *J. Roy. Statist. Soc., A.* (in the press).

² Morison, D. E., and Henkel, R. E. (eds.). *The Significance Test Controversy—A Reader*, (Butterworths, London, 1970).

³ Edwards, A. W. F., *Nature*, **222**, 1233 (1969).

⁴ Edwards, A. W. F., *Likelihood* (Cambridge University Press, in the press).

⁵ Edwards, J. H., *Nature*, **221**, 593 (1969).

⁶ Fisher, R. A., *The Genetical Theory of Natural Selection*, Clarendon Press, Oxford, 1930).

⁷ *Shorter Oxford English Dictionary*, (Clarendon Press, 1959).

⁸ Gray, A., *Cambridge University: An Episodic History*, (Heffer, Cambridge, 1926).

⁹ Fisher, R. A., *Statistical Methods and Scientific Inference*, (Oliver and Boyd, Edinburgh, 1956).

¹⁰ Richards, M., Letter to *The Times* (July 2, 1970).

¹¹ Hogben, L., *Statistical Theory*, (Allen and Unwin, London, 1957).

Fossilmanship in East Africa

SARAH BUNNEY

Recent discoveries in the Plio-Pleistocene deposits of East Africa may call for a re-examination of man's evolutionary history.

HAD a cantankerous camel ridden by Richard Leakey not stalled one August day in 1969 in the desert heat of northern Kenya, palaeoanthropologists might have missed one of the most interesting discoveries of recent years—a skull, intact but for its teeth and jaws, of an early hominid more than two and a half million years old. But for the bad humour of the camel, as Leakey recounted later in the *National Geographic Magazine*¹, he and his party would not have broken camp earlier than was planned and would probably not have come upon the skull lying in a dry stream bed in the Ileret area on the east side of Lake Rudolf. The fossil resembles² *Australopithecus boisei*, the robust form of australopithecine known from East Africa and, by coincidence, Leakey found it almost ten years to the day after his mother, Dr Mary Leakey, discovered the first representative of this species in Olduvai Gorge, farther to the south.

Hunting for fossil man thus has its lucky moments—and the Leakeys certainly seem to have their share of them—but as the pages of *Nature* for the past three or four years show, this activity requires not only a remarkable flare for knowing exactly where to look for fossils but also a willingness to work sometimes under the most inhospitable conditions in what is often unexplored territory; recent palaeontological expeditions to East Africa have made history in more ways than one.

The rewards have, nonetheless, been rich. Richard Leakey's chance discovery in 1969 of the robust austra-

lopithecine skull in the Rudolf area and the associated hominids and fossil vertebrates collected in that year and in the previous year's season were but foretastes of what was to follow in 1970. The results of this last season's work, described earlier this year³, added not only another fifteen or so hominids to the Rudolf list (including what may prove to be specimens of an early representative of *Homo*), but also extended the vertebrate collection to some six hundred specimens⁴.

Richard Leakey is wisely, though tantalizingly, reticent about the taxonomic affinities and dates of these hominids until more detailed anatomical and geological analyses have been made, but there now seems little doubt that this part of East Africa will be as significant in the study of human origins as Olduvai Gorge in Tanzania made famous by Richard's father Dr Louis Leakey.

Being Plio-Pleistocene in age, the Rudolf deposits (see Fig. 1), and those in the Omo Valley in Ethiopia^{5,6}, extend back in time from those at Olduvai by another two million years and, together with the other finds of australopithecines elsewhere in Kenya, they help to fill in the gap between the *Ramapithecus* stage of human evolution, of Miocene times (see the recent reviews of this subject⁷⁻⁹), and the early Pleistocene *Australopithecus* remains from Olduvai and the South African sites. It might therefore be thought that the jigsaw of human evolution was neatly being pieced together. But it is as becoming clear from the preliminary results of the Rudolf expeditions³, not all the pieces are fitting into the puzzle in quite the way that palaeoanthropologists have been led to expect and it seems likely that Richard Leakey's interpretations of the Rudolf material will be as contentious as those his father made regarding the *Homo habilis* specimens from Olduvai.

Coexistence

The chief point of interest in the reports to date^{2,3} of the Rudolf investigations are Richard Leakey's suggestions that more than a million years ago the genus *Homo* coexisted

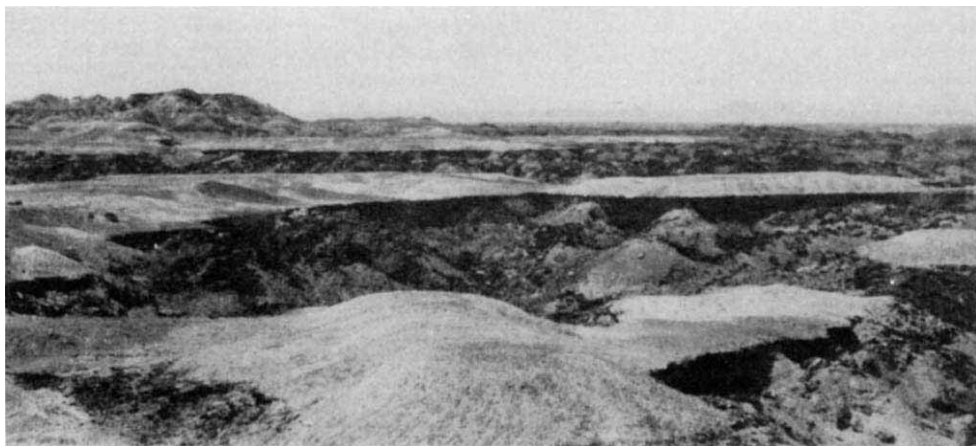


Fig. 1 Pleistocene exposures in the Koobi Fora area, East Rudolf.

with *Australopithecus* in this part of East Africa at the beginning of the Pleistocene, and therefore that the two genera may well have been quite separate and distinct at this time. He also revives the idea that the differences in size and morphology between the gracile and robust types of australopithecine (both of which have been found in stratigraphically similar deposits in the Lake Rudolf area) relate to sexual dimorphism and the two are thus not separate species or subspecies as is usually held to be the case.

Richard Leakey bases his claim for *Homo* at Lake Rudolf on three specimens—two mandibles (both are fragmentary but KNM-ER-730 is more complete than KNM-ER-731) and an incomplete femoral shaft (KNM-ER-737). Further discoveries and more detailed examinations are obviously required before this claim will be seriously considered by other workers—what is now needed, in particular, is some agreed definition of a complex of morphological features characteristic of the line of evolution leading to man—but if these specimens are accepted as *Homo*, and KNM-ER-730 seems a particularly good candidate for inclusion in the genus, they may be the best evidence so far of the genus outside Java where *Homo erectus* is known. Richard Leakey has so far published no dates for his *Homo* sites. It seems very likely that they are more than a million years old, but it is not clear from the reports whether they are as old as the age of 2.61 ± 0.26 million years which is the date given by Leakey² for some of the Rudolf hominids including the specimen (KNM-ER-406), *Australopithecus* aff. *boisei*, mentioned earlier.

This is not the first suggestion that *Homo* and *Australopithecus* were contemporaneous in some parts of Africa. For several years, Louis Leakey has claimed that a more advanced form of hominid, the famous *Homo habilis*, co-existed with *Australopithecus* in the Olduvai Gorge locality for about a million years and that this form and not *Australopithecus* was responsible for fabricating the primitive tools recovered from Beds I and II, known as the Oldowan culture. Only recently¹⁰, he has published an account of another *Homo habilis* specimen (OH 24) which was discovered in 1968 in Bed I at Olduvai. The fossil, nicknamed "Twiggy", was found in a very crushed state and its inclusion in *Homo* is likely to be questioned by the several workers who regard *Homo habilis* as an East African variant of the gracile *Australopithecus africanus*, a species which was known first from the South African sites.

There have been similar doubts regarding the material from Swartkrans in South Africa, once known as "*Telanthropus capensis*" and subsequently reclassified by Professor T. Robinson as *Homo erectus*. According to Robinson and some other workers, this material is evidence of the coexistence in the early Middle Pleistocene of a more advanced hominid than the robust australopithecine. *A. robustus* (formerly *Paranthropus*) known in large number from that site. But this claim has been questioned by some authors who hold that the so called *Homo erectus* material, like *Homo habilis*, is simply another variant of the polytypic *A. africanus* population of the time. This issue was reopened last year with the description by Dr. C. K. Brain and his colleagues¹¹ of a reconstructed skull from Swartkrans which they claim is a representative of the genus *Homo* but to which they assign no specific name. The niceties of this taxonomic problem, like those of *Homo habilis*, will no doubt continue to be argued about by Wolpoff¹² and others for several more years to come.

Part of the reason for the current confusion over early hominid taxonomy lies in the fact that although there is now plenty of cranial material, particularly teeth and lower jaws, postcranial remains of *Australopithecus* are rare. It is only through detailed analyses of such bones that some idea can be obtained of the gait and stance of the creatures. One of the criteria used for distinguishing *Homo* from the

pre-human stage is his upright bipedalism, and the limb bones of early hominids which can provide clues to the origin of bipedal locomotion are therefore particularly valuable. The hominid limb material from Lake Rudolf, claimed by Richard Leakey³ to be representative of both the *Australopithecus* and *Homo* forms, should thus be very useful.

Until the Rudolf discoveries, a complete femoral shaft of an australopithecine was unknown, although femoral ends of robust forms had been recovered several years earlier from Swartkrans (SK 82 and 97) and from Olduvai (OH 20) (see ref. 13 for a discussion of these), and a reconstruction of a femur of a gracile form has been made from material recovered at Sterkfontein. The Rudolf material should therefore allow useful comparisons with this australopithecine material, with the femur of the supposed *Homo* from Rudolf, and also with postcranial material of undisputed *Homo erectus* such as that recovered in 1970 from the Middle Pleistocene Bed IV at Olduvai and described recently in *Nature* by Dr M. H. Day¹⁴. This last material (OH 28) was found in association with an Acheulean industry and is the first direct association of tools with *Homo erectus*. This association naturally does not necessarily imply that *H. erectus* was responsible for the industry, but Mary Leakey¹⁵ thinks it likely because there is other evidence at Olduvai of a possible link between early Acheulean hand axes from Bed II and *Homo erectus*. Day has also been able to compare¹⁴ OH 28 with femora of approximately the same age from Choukoutien in China and with the femora claimed to be of the same age from Trinil in Java.

Antiquity

In spite of the considerable confusion over the fine distinctions between *Australopithecus* and *Homo*, the new work in East Africa is confirming the view that there were marked morphological variations between populations of early hominids and that these variants can be broadly grouped into gracile and robust forms which persisted relatively unchanged for several million years. For just how long they persisted is being brought out by some of the recent discoveries in Kenya and elsewhere. Until a few years ago, australopithecines were thought to have evolved as a distinct group only about a million years ago; even with the re-dating in the early 1960s of the timing of the beginning of the Pleistocene epoch the antiquity of these hominids was thought to be only as much as about 2 to 3 million years. With the discoveries of new Plio-Pleistocene deposits in Kenya and Ethiopia, however, and more reliable radiometric dating techniques it is becoming clear that *Australopithecus* had evolved by the end of the Pliocene.

The oldest undisputed australopithecine material known to date seems to be the portion of a lower jaw with one molar intact found by Professor B. Patterson's team at Lothagam Hill near the south end of Lake Rudolf in Kenya. This fossil is thought to be a representative of the gracile form of australopithecine and its date of 5.5 million years puts it well and truly in the Pliocene. Patterson and his colleagues also have the distinction of finding part of a hominid humerus, which likewise probably belonged to a gracile form in deposits at Kanapoi in Kenya. The date of this specimen was put at 2.5 ± 0.2 million years when it was described in 1967 by Patterson and Professor F. C. Howell, but more accurate dating of the lava flow overlying the deposits in which the fragment was found and comparisons of the mammalian fauna with those from the Omo Valley and elsewhere suggests that the Kanapoi hominid is between 4.0 and 4.5 million years old. In their turn the Omo Valley hominids seems to have a range of between 2.0 and 4.0 million years^{5,6}. At Rudolf the dating has not been completed, but the age already given to the robust skull KNM-ER-406 antedates the earliest hominid

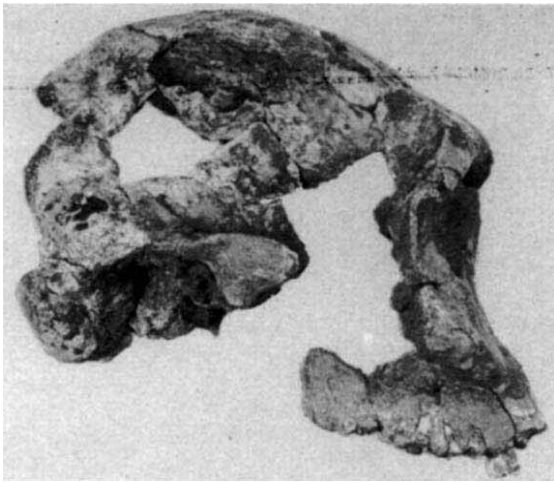


Fig. 2 An almost complete right half of a cranium of an australopithecine collected in 1970 from the Ileret area, East Rudolf.

from Olduvai by about 0.8 million years and similar robust types from Swartkrans and Kromdraai *A. robustus* ($\equiv$ *Paranthropus*), by about 1.5 million years.

Whether australopithecines older than 5.5 million years will be found is another matter, though Bishop *et al.*¹⁶ claim to have recovered the upper molar tooth of a hominid from the 9 to 12 million-year old Ngorora Formation in the Baringo District in Kenya. Louis Leakey has described this specimen as having some affinities to *Kenyapithecus* (equivalent to *Ramapithecus* of Simons and Pilbeam) and some to both *Homo* and *Australopithecus*. Judgment on this find will no doubt be reserved until further discoveries are made from this locality. It is nonetheless clear that the fossiliferous sediments at Baringo are going to be particularly useful in unravelling the evolution of vertebrates during the Pliocene.

It was during the Pliocene that many African mammalian groups had their origin and diversification; other groups, by contrast, became extinct. Comparisons of faunas from different localities can thus provide useful information about the evolution of mammals in Africa and the dating of one faunal assemblage can form a standard of comparison for other faunas. This obviously has a bearing on the correlation of important hominid-bearing sediments. Maglio has already produced a tentative correlation¹⁷ for some Plio-Pleistocene deposits in East Africa based on the remains of fossil elephants found in them. Isolated teeth of Proboscidea are common in these sediments and are often well preserved because of their large size and thick enamel covering and are therefore useful stratigraphical indicators. Maglio has also compared the rich collection of vertebrates from Lake Rudolf⁴ with the assemblages known so far from the Omo Shungura Formation and from the lowermost beds at Olduvai. Two of the species of elephant recovered at East Rudolf include the earliest known occurrences of the *Elephas* and *Loxodont* lineages and indicate that the Koobi Fora I beds at this locality are slightly older than the Kanapoi and Lothagam Hill deposits. There is also an indication¹⁸ at East Rudolf of an older Mio-Pliocene sedimentary and volcanic complex in the Suregei Plateau, and this may provide crucial faunal remains for comparisons with those from Baringo. There is a chance too that some very early hominids will turn up here.

Robust and Gracile

As has been stated, the oldest unequivocal hominid is that dated at 5.5 million years from Lothagam Hill. This

is a gracile form; the first sign of the coexistence of this type with the robust form comes later in time at Omo where Howell and his colleagues⁵ have found fragments of two australopithecine variants which resemble *A. africanus* and *A. boisei* and which are between 3 and 4 million years old. The contemporaneity of robust and gracile forms has subsequently been confirmed by the results of Leakey's three seasons in the Lake Rudolf area. (The question of *Homo* in these deposits was discussed earlier; whether it is *Homo* or a variant of the gracile form does not alter the concept of two coexisting populations of early hominids.)

There has, of course, been much discussion about the relationships of the types of robust australopithecine—the northern *A. boisei* and the southern *A. robustus*. It is held by some workers, for example Tobias¹⁹, that the East African robust form was ancestral to *A. robustus* and that the two were members of an over specialized line of early hominids which became extinct in the Middle Pliocene. The genus *Homo*, according to this theory, evolved from the less specialized gracile *Australopithecus*. It has also been suggested that the size and morphological differences between the gracile and robust types relate to sexual dimorphism, but this view has not been widely accepted because it does not explain why at some sites there is a predominance of one type only and it is considered unlikely that just one sex, if the australopithecines were indeed widely sexually dimorphic, would be preserved at these sites. The recovery by Richard Leakey and his team of both types from the same stratigraphic levels should reopen this thorny problem. As already mentioned, Leakey believes that the two types represent the two sexes of the same australopithecine species and that this population need not necessarily have been in competition for resources with his supposed *Homo* population; *Homo* and not *Australopithecus* would have been responsible for the Oldowan culture (see Fig. 2) recovered in 1969 from the Koobi Fora area (with a date of 2.6 ± 0.26

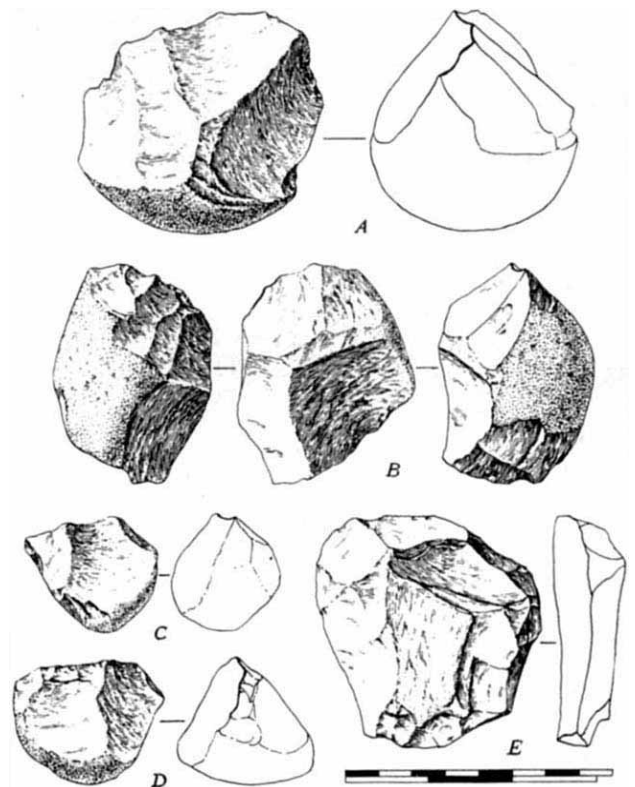


Fig. 3 Oldowan culture collected in 1969 from the Koobi Fora area, East Rudolf.

million years, it is the oldest collection of unequivocal artefacts known).

If *Homo* and *Australopithecus* did share a habitat at Rudolf in the early Pleistocene and the gracile form was indeed the female of the robust form, long established ideas in the palaeoanthropological world really would be upset. It would certainly require the re-examination of the theory that the gracile australopithecine was ancestral to *Homo*, and raises the interesting possibility that all the early Pleistocene australopithecines and not just the robust forms were members of a side branch of human evolution and doomed to extinction in the Middle Pleistocene.

The suggestion that *Homo* and *Australopithecus* were evolving in parallel in East Africa between three and one million years ago is not likely to be generally accepted at present, but it is one that deserves serious discussion. In this connexion, the recently described robust australopithecine cranium dated at between 1.1 and 1.2 million years from Chesowanja in the Baringo District is interesting. This specimen is described by Carney *et al.*²⁰ as less specialized than either the *A. robustus* or *A. boisei* types, and they claim that it possesses certain features characteristic of *Homo*. The cranium, unfortunately, is badly preserved and its significance in the evolution of man is questionable.

It is to be hoped that more important finds will turn up in the next few years. Who knows what might be recovered in the Lake Rudolf and Baringo area. Has the early hominid ancestor of *Homo* yet to be found?

- ¹ Leakey, R. E. F., *Nat. Geog. Mag.*, **137**, 712 (1970).
- ² Leakey, R. E. F., *Nature*, **226**, 223 (1970).
- ³ Leakey, R. E. F., *Nature*, **231**, 241 (1971).
- ⁴ Maglio, V. J., *Nature*, **231**, 248 (1971).
- ⁵ Howell, F. C., *Nature*, **223**, 1234 (1969).
- ⁶ Brown, F. H., and Lajoie, K. R., *Nature*, **229**, 483 (1971).
- ⁷ *Nature*, **219**, 1820 (1968).
- ⁸ Tattersall, I., *Man's Ancestors: An Introduction to Primate and Human Evolution* (Murray, London, 1970).
- ⁹ Pilbeam, D., *The Evolution of Man* (Thames and Hudson, London, 1970).
- ¹⁰ Leakey, M. C., Clarke, R. J. and Leakey, L. S. B., *Nature*, **232**, 308 (1971).
- ¹¹ Brian, C. K., *et al.*, *Nature*, **225**, 1112 and 117 (1970).
- ¹² Wolpoff, M. H., *Nature*, **230**, 398 (1971).
- ¹³ Day, M. H., *Nature*, **221**, 230 (1969).
- ¹⁴ Day, M. H., *Nature*, **232**, 383 (1971).
- ¹⁵ Leakey, M. C., *Nature*, **232**, 380 (1971).
- ¹⁶ Bishop, W. W. and Chapman, G. R., *Nature*, **226**, 914 (1970).
- ¹⁷ Maglio, V. J., *Nature*, **225**, 328 (1970).
- ¹⁸ Vondra, C. F., *et al.*, *Nature*, **231**, 245 (1971).
- ¹⁹ Tobias, P. V., *Olduvai Gorge*, **2**, (Cambridge University Press, (1967).
- ²⁰ Carney, J., *et al.*, *Nature*, **230**, 509 (1971).

Problems of Artificial Fertilization

R. G. EDWARDS

UNTIL recently, opportunities for scientific and clinical studies on human conception have been extremely limited, largely because of the inaccessibility of the ovary and the difficulty in obtaining preovulatory oocytes. One egg per month was a discouraging target and experience had shown that even this single egg was difficult to recover because ovulation could not be predicted with any degree of accuracy. In the past two or three years, considerable progress has been made with the clinical and scientific problems associated with the study of human fertilization and cleavage. The opening came with the discovery that human oocytes would undergo their final stages of maturation—those occurring just before ovulation—when removed from their follicles and placed in a suitable culture medium. These stages could be timed with considerable accuracy, and the availability of this human material provided opportunities for studies on meiosis, ovulation and fertilization¹. Oocytes matured *in vitro* thus provided the material for studies on human fertilization *in vitro*, which was achieved two years ago.

Patients Involved

When fertilization had been accomplished using these oocytes, the emphasis of the work changed to include studies on patients in order to realize the potential of these studies for human infertility. There were various reasons for this extension into clinical studies. While oocytes matured *in vitro* provide good material for fertilization, they are unsuitable for studies on cleavage because observations in animals has shown that the resulting embryos fail to develop normally². For normal development to full term, oocytes must be collected after completing, or almost completing, their maturation in the ovary (see Fig. 1). Endocrine methods were thus needed to induce follicular growth and oocyte maturation in the patients, the preovulatory growth of oocytes had to be timed exactly from the estimates made from cultures and surgical methods had to be developed for the collection of the oocytes just before ovulation. The achievement of these objectives has opened the possibility

of growing and analysing cleaving embryos in culture, replacing them in the mother, and studying the processes of implantation. Each of these projects opens fundamental opportunities for the study of human conception. What progress has been made in these various areas of investigation?

The surgical methods for collecting human oocytes are perhaps almost as refined by now as they ever will be. Laparoscopy has proved to be a simple and very safe procedure permitting the necessary manipulations to be carried out on the ovary³. Oocytes are collected from more than one-half of the available Graafian follicles, and although the proportion could be improved the loss of one-half of the oocytes in this way is not holding up further developments in any serious manner. The timing of oocyte maturation and ovulation are also under fine control. None of the many patients so far examined has ovulated before the expected time, and many of the oocytes were fertilized soon after

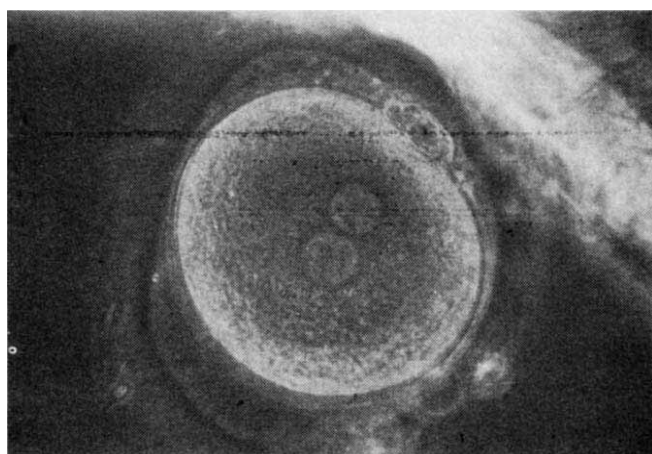


Fig. 1 Human pronucleate stage.

collection, showing that they were certainly preovulatory. The major problems in the initiation of follicular growth and oocyte maturation concern the response of the patients to the endocrine treatments, which are based on the use of human menopausal gonadotrophins on amenorrhoeic women⁴. Some of the patients have few or no preovulatory follicles after the treatment.

There are well known difficulties with the method—for example, the great variation between patients or even between successive treatments of the same patient⁵. Three or four preovulatory follicles are the minimum required to provide sufficient oocytes to work with. Measurements of the excretion of oestrogen in the urine of the patients during treatment provides some indication of their response to the treatment, particularly as a safeguard against overstimulation, but its usefulness as a measure of the growth of preovulatory follicles remains to be decided. There is some scope for improvements in the hormone treatment. The amount of gonadotrophin being administered to the patients is modest in comparison with that use for the treatment of amenorrhoea, partly because the patients have a normal menstrual cycle which might heighten their response. Simply raising the dosage of gonadotrophins should induce a greater response in most patients. Alternatively, the use of compounds such as clomiphene, to induce the release of endogenous hormones, might prove superior, cheaper, and more acceptable than the injection of gonadotrophins. Clomiphene is widely used in the treatment of female hypopituitarism, and should work well with cyclic patients. The recent purification of the gonadotrophin releasing hormones of the hypothalamus could well offer another alternative, provided that the dosage and response can be controlled.

Human Ova

Studies on the fertilization of mammalian ova *in vitro* have moved quickly in the past two or three years after a decade of uncertainty. Belief in the necessity of uterine spermatozoa had become almost dogmatic as a result of experiments involving mostly rabbit ova. Even the demonstration that hamster ova could be fertilized *in vitro* by epididymal spermatozoa⁶ had scarcely shaken this fixed belief. When the conditions needed in the culture medium for hamster fertilization were defined, the work was extended to human material, using oocytes matured *in vitro* and washed, ejaculated spermatozoa. Fertilization was established using the same criteria defined for studies with animal ova⁷. Moreover, similar studies are now being extended to other species such as the mouse and cow, and even in rabbits spermatozoa have now been found to penetrate through the membranes of the ova. The culture media are simple: a basic Tyrode's solution is supplemented with bovine serum albumin, pyruvate, and bicarbonate to a pH of 7.6. Pyruvate is obviously an energy source, and the albumin presumably serves as a source of nitrogen or to stabilize cell membranes. The significance of the high pH, or of the contributions of the cumulus cells surrounding the ova, remain unexplained.

Media suitable for fertilization and for cleavage are rarely identical. The development of media to support the cleavage of human ova has not relied on experiments with mammalian ova to nearly the same degree. Simple defined media had been developed for the cleavage of mouse ova, following the recognition of the importance of pyruvate as an energy source during these early embryonic stages⁸. The initial attempts to culture human embryos utilized these media, but cleavage progressed only to the 8-celled stage⁹ (Fig. 2). More complex media developed for routine tissue culture have proved much more successful. A medium composed of Ham's F 10 supplemented with calf or human serum, at a pH of 7.3 under a gas phase of 5% O₂, 5% CO₂ and 90% N₂, has supported development to the blastocyst stage¹⁰. Several

blastocysts, each having 100 or more nuclei and several mitoses, have been grown under these conditions (Fig. 3). The inner cell mass, trophoblast and blastocoelic cavity were well defined, and some blastocysts expanded in culture. The blastocoel appeared to form from the secretions of large cells that accumulated at one pole of the morula. Modified body fluids are also reported to support the cleavage of human embryos¹¹. According to one report, oocytes matured *in vitro* had to be fertilized by human spermatozoa previously incubated in the uterus of a monkey¹²; the abortive cleavage of the embryos might have been a consequence of maturation *in vitro* or some other cause.

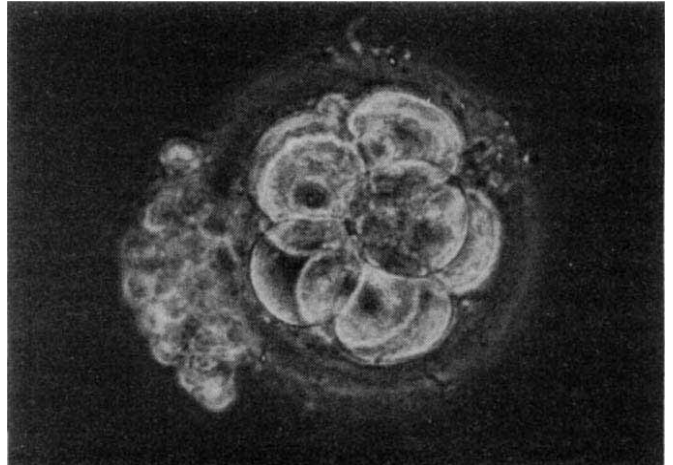


Fig. 2 Human eight-celled embryo.

Where Next?

Further development beyond the blastocyst stage must surely be the next study. This raises the question of whether the embryos will be able to implant in the mother, that is, if the uterus has been sufficiently stimulated by the endocrine treatments used to induce follicular maturation. Another question concerns the capacity of the embryos to implant. The indications are good: the pronounced differentiation of many embryos in culture augurs well for their continued development. Conditions in the mother also seem to be satisfactory: most patients have had a secretory endometrium, excreted suitable amounts of pregnanediol, and their induced cycle has resembled a natural cycle³.

What is known of the capacity of the embryos for normal development? The criteria for cleavage are now well recognized, and provide the first safeguard. Several of the embryos have possessed a chromosome complement that appeared to be diploid, thus excluding the possibility of triploidy. Trisomy has not been excluded. Yet the more that

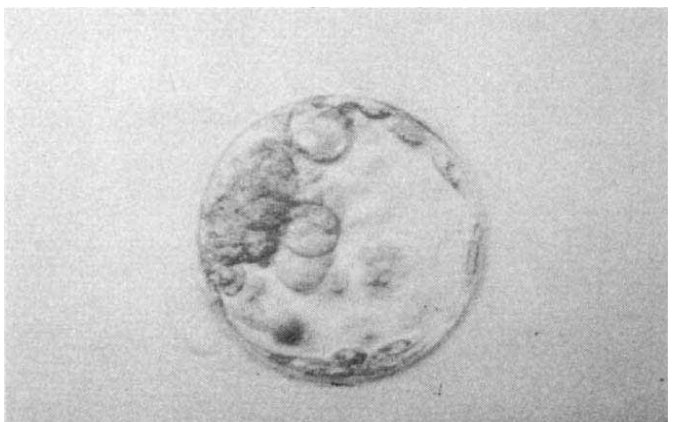


Fig. 3 Human blastocyst.

is learnt about natural human pregnancy, the more obvious becomes the conclusion that many embryos conceived naturally are trisomics, triploids or mosaics destined to perish in the first trimester. The possibility of trisomy in cultured embryos thus becomes less important, and especially since Down's syndrome can be identified in foetuses by amniocentesis. Nor should the problem of teratogenic development present undue difficulties, because widespread experience in animals and man has shown that agents known to be teratogenic in later stages of pregnancy are lethal to preimplantation embryos.

The immediate social consequences of the successful re-implantation of human embryos would be the cure of some forms of infertility, notably where the wife has occluded oviducts or the husband is oligospermic. There should be no criticism in giving these couples their own children: comments about overpopulation seem to be highly unjust to such an underprivileged minority.

The next development could well be the sexing of embryos before transfer. Rabbit blastocysts were sexed from the presence or absence of sex chromatin in small pieces of trophoblast excised by microsurgery¹³; the embryos recovered after a few hours in culture and some of them developed to full term in recipient females. The sex of all foetuses had been predicted correctly; this was the first occasion that the secondary sex ratio had been controlled in any mammal. Most efforts in this direction had been devoted to the attempted separation of X and Y spermatozoa, so far without any convincing success.

Improvements are now needed in the methods for sexing blastocysts. Two approaches are currently feasible: the identification of enzyme activity or of antigens determined by sex-linked genes. The activity of glucose-6-phosphate dehydrogenase (G6PD) and hypoxanthine-guanine phosphoribosyl transferase (HGPRT) have been studied in mouse embryos; it is not yet certain that these genes are sex-linked in mice as they are in man. Assuming that they are sex-linked, female embryos might be expected to possess twice the activity of male embryos. Unfortunately, levels of G6PD were dominated by the massive amounts of maternal enzyme "inherited" from the oocyte, and which declined steadily during cleavage^{8,14}. The activity of HGPRT in embryos increased after fertilization, indicating that this enzyme was being synthesized during these early stages and, perhaps more important, that the sex chromosomes are active at this time¹⁵. This enzyme might thus be present in different

amounts in male and female embryos. Evidence for sex-linked antigens is restricted to one report showing that the sex ratio of mouse offspring can be altered by interfering with the normal immunological processes of pregnancy¹⁶. Sexing blastocysts would be an excellent approach to the control of sex-linked mutant genes in man by avoiding the birth of affected boys. Strict control would have to be exerted over sexing if it could be demonstrated that widespread parental choice would seriously disturb the sex ratio in human populations.

Studies on human conception open new approaches to human problems including preventive medicine and family planning. More knowledge about the growth of ovulatory follicles, the secretions of the blastocyst or the components of capacitation could lead to the development of novel methods of contraception. Rational discussion is needed about the emerging social issues involved in these studies. Much of the debate so far has been unreal in the sense of being far removed from the actual clinical problems involved: constant reference to nuclear cloning is an example. Earlier remonstrations about the disposal of tiny blastocysts appear almost irrelevant when foetuses of four months' gestation or older are being aborted for social reasons. Alleviating infertility has even been criticized because more consideration is needed about the survival of the family unit in our society! Realization of the benefits that could accrue to many people should lead to more rational discussion and conclusions.

¹ Edwards, R. G., *Lancet*, ii, 926 (1965).

² Chang, M. C., *J. Exp. Zool.*, **128**, 379 (1955).

³ Steptoe, P. C., and Edwards, R. G., *Lancet*, i, 683 (1970).

⁴ Hack, M., Brish, M., Serr, D. M., Insler, V., and Lunenfeld, B., *J. Amer. Med. Assoc.*, **211**, 791 (1970).

⁵ Crooke, A. C., *Brit. Med. Bull.*, **26**, 17 (1970).

⁶ Yanagimachi, R., and Chang, M. C., *J. Exp. Zool.*, **156**, 361 (1970).

⁷ Edwards, R. G., Bavister, B. D., and Steptoe, P. C., *Nature*, **221**, 632 (1969).

⁸ Brinster, R. L., Chap 17 in *Biology of the Blastocyst* (edit. by Blandau R. J.) (University of Chicago Press, 1971).

⁹ Edwards, R. G., Steptoe, P. C., and Purdy, J. M., *Nature*, **227**, 1307 (1970).

¹⁰ Steptoe, P. C., Edwards, R. G., and Purdy, J. M., *Nature*, **229**, 132 (1971).

¹¹ Shettles, L. B., *Nature*, **229**, 343 (1971).

¹² Gardner, R. L., and Edwards, R. G., *Nature*, **218**, 346 (1968).

¹³ Brinster, R. L., *Biochem. J.*, **101**, 161 (1966).

¹⁴ Epstein, C. J., *J. Biol. Chem.*, **245**, 3289 (1970).

¹⁵ Lappe, O. O., and Schalk, M. J., *Transplantation*, **11**, 491 (1971).

Assembly of Tobacco Mosaic Virus Particle

P. J. G. BUTLER

MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH

TOBACCO mosaic virus (TMV) is one of the best characterized viruses and consists of a long rod, which is composed of a helical array of identical protein subunits with the single stranded RNA lying in the helical groove between adjacent turns of the protein helix (for review see ref. 1). TMV was the first virus to be reconstituted from its RNA and protein¹ and has frequently been used as an example of the self-assembly process, and the protein alone is capable of forming rods which are structurally indistinguishable from the virus helix (see ref. 2). However, the protein also has many other polymorphic forms and this property complicates the understanding of the protein aggregation and has led to different theories being suggested to account for this process^{2,3}.

Recently, in a concerted attack on this problem in the MRC Laboratory in Cambridge, the structures of these different aggregates have been determined, using the electron microscope and X-ray diffraction techniques, and the conditions in which they occur defined⁴. The results support the ideas put forward some years ago by Caspar² that all the aggregates are the result of the same specific interactions as are found in the virus but slightly modified in the different polymorphic forms. The results also allow the mechanism controlling which protein aggregate will occur to be determined⁵. One of these aggregates, the disk, is found only in the conditions in which virus assembly occurs and it has been shown that growth of the virus does not occur by the addition of single protein molecules but rather by the in-

corporation of a whole disk at once⁶. The polymorphism of TMV protein is thus the basis for the rapid and specific formation of the virus from its protein and RNA.

Protein Aggregates

The aggregation of TMV protein can occur in one of two modes. The first to be structurally understood was the helix, in which the arrangement and conformation of the protein subunits is identical to that in the virus. The aggregation of the protein in this mode can be described as "quasi-crystallization" and the resulting rods are of indeterminate length, usually very large. The second mode can be described as "two-layer" aggregation and aggregates of this type are stabilized by an additional "pairing" interaction, which is not present in the helix, between the subunits vertically above each other. The movements of about 5 Å of the outer ends of the subunits, involved in this pairing, have been seen by 3-D image reconstruction from electron micrographs⁷. Because of the detailed geometry of the TMV protein subunit and of the intersubunit interactions, the layers curve round and the ends meet to form a closed torus consisting of two rings, each of 17 subunits (Fig. 1). It is this structure which is called the "disk".

The growth of the two-layer aggregates, from the trimer to the 32-mer, can be described as linear condensation polymerization, with each subunit adding with the same binding energy. The formation of the trimer, which begins the association process, and the addition of the final two subunits to close the disk cannot be described so simply. However, the energies of these various interactions have been calculated⁵ and the resulting quantitative model does predict the observed sedimentation patterns.

A further complication with the two-layer aggregates is that under favourable conditions for the bonding interactions (at high ionic strength) the disks can stack on each other, to give even larger aggregates. This can result in a series of discrete, large aggregates in equilibrium with a range of small aggregates and gives rise to the sedimentation patterns which have caused much argument in the past. Still further complication is caused by the readiness with which the aggregation of TMV protein can become irreversible, presumably by some slight damage to the protein, and hence give aggregates which are no longer in equilibrium with the rest of the mixture. One such structure, the stacked disk rod, consisting of long rods of disks stacked on each other in a polar fashion, has recently proved important in the elucidation of both the structure^{4,7} and control of formation⁵ of the two-layer aggregates, just because of its size and the irreversibility of the disk aggregation within it, but even this had earlier caused considerable difficulties in the interpretation of experiments in which it had occurred.

The bonding in these two modes is very similar, with the pairing perturbation in the two-layer aggregates as the only major difference⁴. The importance of this pairing in limiting the growth of the aggregates to a single direction and thus causing the formation of a discrete aggregate, the disk, rather than the indeterminate sized helix, lies in the assembly process for the viral rods, which I shall discuss later.

When the conditions in which any particular aggregate contributes significantly to the equilibrium are determined (Fig. 1 of ref. 4), the pH is found to be the controlling variable for the mode of aggregation. As hydrophobic bonding is the driving force for aggregation³, increasing both the temperature and the ionic strength will favour aggregation and cause large aggregates to form, but neither has a major effect on the mode of aggregation. The exact pH at which the transition between the modes occurs depends slightly on the temperature and ionic strength, but it is always about pH 6 to 7, with the helical aggregation predominating below the transition and the two-layer aggregates above. This transition pH is found to correspond to the pKs of two carboxyl groups, which titrate abnormally in the virus² but not in the stacked disk rods⁵, and in the titration of the protein helix an abnormally sharp release or uptake of protons correlates with the disaggregation or formation of the helix^{8,9}.

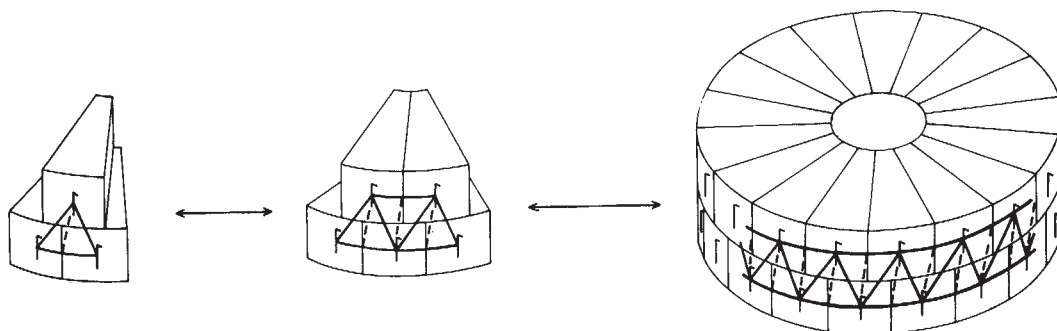
Control of Aggregation

The control effect can be seen qualitatively when the pH of a solution of the protein helix is raised and the turbidity, due to the long rods, disappears immediately as the rods dissociate to give the small protein aggregates (called the "A-protein"). If the pH is lowered again soon afterwards, the turbidity returns as the protein subunits reform helix. An important variation occurs if the protein is kept above the transition pH, in conditions where it forms disks, for long enough for the disks to be formed before the pH is lowered. In this case short, imperfect, helical rods are formed within 5 seconds, and these then anneal over a period of about a day to give long, perfect helices again. Close examination of the imperfect helices shows that they are composed of two turn segments of helix stacked on top of each other, with their ends not in register⁴, each segment appearing to be formed from a single disk by it having dislocated under the influence of the high proton concentration.

Since the subunits in any single disk must change their mode of aggregation together, any such transition is inherently cooperative and can be considered to be allosteric, with the protons as the allosteric effectors. In this case the effectors act by protonating two carboxyl groups, to give two carboxyl-carboxylate pairs. This abolishes the repulsion between the charged carboxyl groups and concomitantly the pairing distortion of the two-layer aggregate, enabling the subunits to take up the strictly equivalent positions of the helix. Were it not for the carboxyl-carboxylate pairs the bonding in the helix would presumably be more favourable than that in the disk, but the unfavourable interaction of these groups acts as a negative switch to prevent helix being the most favoured aggregate under all conditions.

This negative, and energetically unfavourable, switching is in contrast to the position that is found in the cases where ionic interactions have been investigated in detail in other proteins. In haemoglobin, which in many ways forms a close parallel with the situation in TMV protein, salt bridges are formed between the chains in the deoxy form but not

Fig. 1 Model for two-layer aggregation to form disks⁵. Subunits associate laterally in a linear fashion with polymerization in the axial direction limited by the pairing deformation⁴ (shown by the dotted line). Other intersubunit contacts are shown by solid lines.



in the oxy form¹⁰. These bridges have to be broken to allow the quaternary structure change, which also involves a movement between the chains of about 5 Å, brought about by the ligand, to go from the deoxy to the oxy structure^{10,11}. In both haemoglobin and TMV protein the allosteric transition is also influenced by other ligands, diphosphoglycerate in the case of haemoglobin and with TMV protein the viral RNA.

Protein and RNA

TMV has 16½ protein subunits per turn of the helix and one problem in its assembly is the nucleation of the helix, as a single ring of subunits joined side to side would be readily broken before it had grown to form a small helix. With the growth of a two-layer aggregate (Fig. 1) this problem would not occur as a subunit in either layer could bond to two subunits in the other layer as well as to its neighbours in the same layer, to give a relatively strong, cross-strutted structure. It has also been seen, from the effect of the pH drop, that a disk is capable of being transformed into a small helix. In the relevant biological process this is brought about by interaction with the viral RNA rather than by protonation⁶. This interaction occurs between a special sequence of about 50 nucleotides at the 5'-hydroxyl end of the TMV RNA (ref. 6) and two protein disks⁹, possibly as shown in Fig. 2a. After the disks have formed short helices, entrapping the first turn of the RNA (Fig. 2b), the initiation is complete and growth can ensue. In general, growth will occur with the RNA protruding from one end of the growing nucleoprotein helix, and the first step will be the interaction of this with a further protein disk which dislocates to become helical and "unrolls" onto the top of the nucleoprotein helix

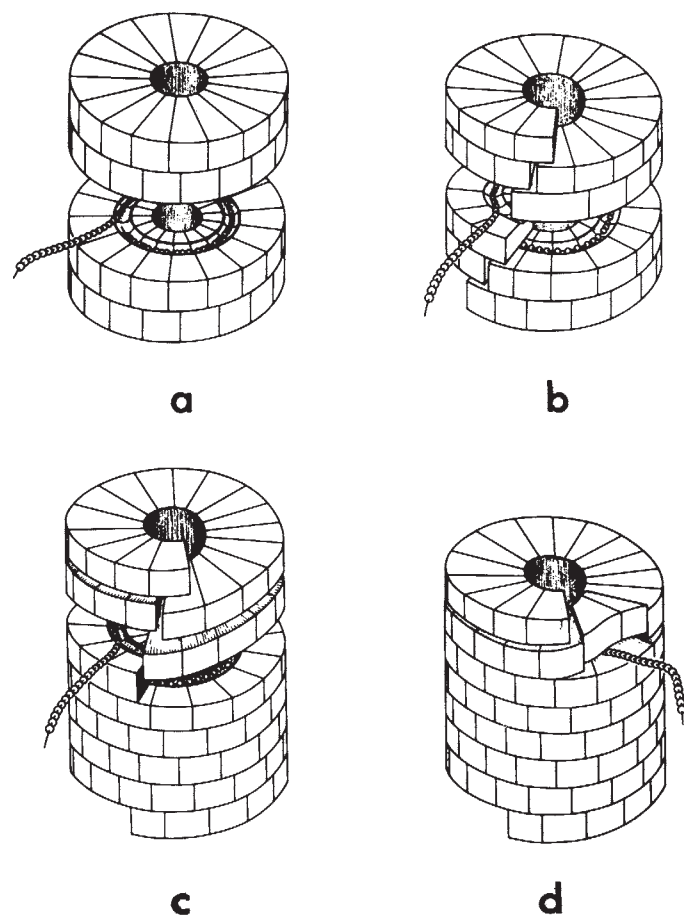


Fig. 2 Model for TMV assembly from disks and RNA⁹. Initiation—*a*: 5'-hydroxyl end of TMV RNA interacts with two disks; *b*: disks dislocate to form short helices enclosing end of RNA. Growth—*c*: dislocated disk "unrolls" to entrain turn of RNA; *d*: second turn of RNA intercalates through "ruck" formed at end of first turn of helix.

to enclose one turn of the RNA (Fig. 2c). A further turn of the RNA can be enclosed by the second turn of the protein helix by intercalation through the "ruck" which is formed at the end of the first turn, where the RNA protrudes sideways (Fig. 2d).

Because of the special nature of the initiation reaction it is very fast and the subsequent growth is relatively slow, unlike most polymerization reactions where the initiation is the slow step. The rate limiting step of the TMV assembly is the actual rearrangement of the protein subunits from the disk into the nucleoprotein helix and is fast enough for the whole assembly to be complete within a few minutes.

General Remarks

The possible role of the disks in nucleation of TMV assembly was foreseen⁶ and, indeed, provided the motive for our early experiments with disks and the RNA, but although in hindsight the continued growth with disks as the protein source has several advantages, it was completely unexpected since everyone expected that the virus would grow by adding single subunits at the end of the growing helix, rather like crystal growth along a screw dislocation. The problem of helix initiation has already been discussed and the biological solution for this, using two disks, confers a high degree of selectivity for its homologous RNA onto the protein. This comes about because each protein subunit interacts with three nucleotides in the RNA, and hence 51 nucleotides in all are involved in a single turn of the nucleoprotein helix. In the initiation step these nucleotides bring about the transition of two disks into helix, involving 68 protein subunits, and this apparently requires an especially favourable interaction. By contrast, the use of disks for the further growth will allow even unfavourable nucleotide sequences to interact and to be coated, as the cooperative transition of the protein subunits will result in the favourable interaction nearby overcoming the local unfavourable effect.

Although the formation of disks is slow, taking several hours at normal protein concentrations, these could be formed before the RNA is synthesized and then be available to coat and protect it very rapidly in the plant cell. The optimum conditions for the formation of disks (pH 7, ionic strength about 0.1) are about those which might be expected in the cytoplasm of a tobacco plant cell, as must be the case if this mechanism is the true biological one.

Clearly it would be a disaster for the virus if its protein alone could form a stable structure over a wide range of conditions, and so not be available to interact with the viral RNA. The role of the polymorphism of the protein is to prevent the occurrence of this disaster, so that instead of being an adventitious and bothersome feature it is now seen to be the controlling factor for the rapid and successful growth of the virus. This polymorphism is itself controlled by a negative switch which prevents the very stable helix being the dominant form of the protein alone over all moderate conditions, but which can be switched on by the viral RNA to give just this stable product as the virus.

¹ Caspar, D. L. D., *Adv. Prot. Chem.*, **18**, 37 (1963).

² Frankel-Conrat, H., and Williams, R. C., *Proc. US Nat. Acad. Sci.*, **41**, 690 (1955).

³ Lauffer, M. A., and Stevens, C. L., *Adv. Virus Res.*, **13**, 1 (1968).

⁴ Durham, A. C. H., Finch, J. T., and Klug, A., *Nature New Biology*, **229**, 37 (1971).

⁵ Durham, A. C. H., and Klug, A., *Nature New Biology*, **229**, 42 (1971).

⁶ Butler, P. J. G., and Klug, A., *Nature New Biology*, **229**, 47 (1971).

⁷ Finch, J. T., and Klug, A., *Phil. Trans. Roy. Soc., B*, **261**, 211 (1971).

⁸ Klug, A., and Durham, A. C. H., *Cold Spring Harbor Symp.*, **36** (1971, in the press).

⁹ Butler, P. J. G., *Cold Spring Harbor Symp.*, **36** (1971, in the press).

¹⁰ Perutz, M. F., *Nature*, **228**, 726 (1970).

¹¹ Perutz, M. F., *Cold Spring Harbor Symp.*, **36** (1971, in the press).

Vintage Year for Tumour Virology

JOHN TOOZE

REVIEWING the results of a year's work in an area of research as active as tumour virology is an invidious as well as a difficult task, for with such a wealth of new information choice is nothing short of embarrassing and sins of omission are all too easy to commit. But no one would, I believe, argue against giving pride of place to reverse transcriptase, the enzyme which Temin and Mizutani and Baltimore (*Nature*, **226**, 1209, 1211; 1970) simultaneously discovered in the particles of four animal and avian tumour viruses, for the significance of the discovery of this enzyme extends far beyond tumour virology. Reverse transcriptase is, of course, the enzyme which can use the single stranded RNA genomes of RNA tumour viruses as templates for the synthesis of first an RNA-DNA hybrid and then a double stranded DNA.

With so many research groups now aboard the bandwagon, it is hard to believe that the enzyme was discovered less than eighteen months ago. It is perhaps even harder to believe that its existence was predicted by Temin as long ago as 1964, when he speculated about the way in which RNA tumour viruses transform cells which have DNA genomes. Temin argued that to become stably associated with the genome of a cell, and to maintain the cell in a transformed or malignant state, the RNA genome of a tumour virus must be transcribed into a double stranded DNA molecule, which he called the provirus. Once the DNA provirus carrying the genes which cause transformation had been made, it could integrate into the cellular genome and be inherited along with the cellular chromosomes. Indirect evidence suggested that transformation by RNA tumour viruses depends on DNA metabolism rather than RNA metabolism and the discovery of reverse transcriptase, an enzyme with the necessary capability, conclusively vindicated Temin's heterodox ideas; it also, of course, provided countless of his colleagues with new research programmes.

Within weeks of the announcement of the discovery of reverse transcriptase, Spiegelman and his colleagues had screened as many different sorts of tumour viruses as they could lay their hands on and had found the activity in all of them (*Nature*, **227**, 563; 1970); today the tally exceeds thirty. Moreover, two viruses—visna virus, which causes a slowly progressing neurological disorder in sheep, and progressive pneumonia virus, which causes a slowly evolving disease of sheep lung have been shown not only to contain reverse transcriptase but also to possess previously unsuspected oncogenic potential; they transform cells *in vitro* (Takemoto and Stone, *J. Virol.*, **7**, 770; 1971). In short, reverse transcriptase as Temin and Baltimore defined the enzyme, that is an enzyme which can use natural single stranded RNA as a template for the synthesis of double stranded DNA, has been found in every RNA tumour virus so far examined.

Tumour virus genomes are not the simplest of RNAs to work with and not surprisingly, Temin (*Nature*, **228**, 424; 1970) and Spiegelman (*Nature*, **228**, 430; 1970) and their colleagues, and other groups, soon began using as templates for reverse transcriptase synthetic RNAs, DNAs and hybrids of RNA and DNA which have a simple and known base composition. Some of these synthetic templates were discovered to have another very great advantage; they were used by the enzyme more efficiently, by almost two orders of magnitude, than the viral genome. For a few weeks if not months this finding

caused considerable excitement, for with such very efficient templates for reverse transcriptase at hand it seemed reasonable to hope to detect the enzyme in human cancer cells, supposing they had been transformed by putative human cancer viruses. And sure enough, several groups, for example Gallo and his colleagues (*Nature*, **228**, 927; 1970), found enzyme activities in human leukaemic cells which could use synthetic nucleic acids as templates for the synthesis of DNA; transiently the way seemed open to tracking down human cancer viruses by following the reverse transcriptase footprints they had so obligingly left behind them.

But the bubble soon burst. For although human cancer cells undoubtedly contain a DNA polymerase activity which can be programmed by synthetic double stranded RNAs and hybrids of RNA and DNA, it soon became apparent that such activities are by no means restricted to cancer cells; they are present in normal cells. Todaro and his colleagues, for example, have not only shown that (*Nature*, **229**, 318; 1971), but, using mouse 3T3 cells infected with mouse leukaemia virus, they have separated the reverse transcriptase specified by the virus from the cellular enzyme which can be programmed by synthetic hybrids to make DNA of RNA and DNA (*Nature New Biology*, **231**, 163; 1971). The salutary lesson of this episode was simply that synthetic nucleic acids are not specific templates for reverse transcriptase; they are most useful for characterizing the enzyme once its existence in a cell or virus has been proven in experiments with its natural RNA template, but synthetic nucleic acids cannot be used to establish the enzyme's existence.

Human Cancer Virus

This particular exploitation of reverse transcriptase in the search for human cancer viruses thus ended abruptly in disillusionment. More recently, however, Moore and his colleagues (*Nature*, **229**, 611; 1971) have detected a virus in samples of human milk which in structure closely resembles mouse mammary tumour virus. This murine virus causes mammary cancer in mice and its existence inspired Moore *et al.* to search for a human counterpart. The human milk virus they found seems to be more prevalent in families or populations with a particularly high incidence of breast cancer; moreover, it contains reverse transcriptase which can use as template the endogenous single stranded RNA genome, as Schlom *et al.* have shown (*Nature*, **231**, 97; 1971). Human milk virus is, therefore, the most likely candidate for the distinction of being the first human cancer virus to be identified.

The central role of a reverse transcriptase activity in Temin's provirus hypothesis has been touched on briefly (see above). But what experimental evidence is there to prove that the reverse transcriptase which everyone is assaying in animal tumour virus particles plays any role in the process of transformation? As more information about the properties of this enzyme and the drugs which inhibit it is being accumulated a trickle of circumstantial evidence which implicates reverse transcriptase in transformation is appearing. Streptovaricins, for example, which block reverse transcriptase also, according to Brockman and his colleagues, prevent the transformation of mouse fibroblasts by mouse sarcoma virus. Furthermore, streptovaricins

delay the induction of leukaemia in mice by murine leukaemia virus. Such indirect evidence, however, can always be interpreted in more than one way and the experiments Hanafusa's group reported (*Virology*, **43**, 313; 1971) provide more convincing evidence of the role of reverse transcriptase in transformation. Before the enzyme had been discovered, his group had isolated a variant of Rous sarcoma virus which appeared to be completely uninfected and, unlike wild type virus, failed to transform chick fibroblasts. After the discovery of reverse transcriptase Hanafusa *et al.* tested their variant and found it lacked activity. The obvious conclusion is that the variant carries a mutation which inactivates reverse transcriptase and renders the virus uninfected.

To bolster this interpretation of the properties of their variant, they infected chick fibroblasts with the uninfected sarcoma virus and an avian leukaemia virus which can replicate in these cells but does not transform them. After such double infections the fibroblasts were transformed and yielded progeny sarcoma virus capable of transforming other fibroblasts. However, the fibroblasts transformed by these progeny sarcoma viruses yielded not infectious progeny virus but uninfected virus identical to the original variant. The most plausible explanation of these results is as follows: the variant lacks reverse transcriptase and as a result is uninfected and unable to transform; in the presence of an avian leukaemia virus the sarcoma virus genome can use the reverse transcriptase specified by the leukaemia virus and can as a result transform cells; the progeny sarcoma virus particles produced by cells infected with these two viruses are infectious and transforming because they can pick up reverse transcriptase specified by the leukaemia virus; the cells transformed by the progeny sarcoma viruses carrying the leukaemia virus enzyme cannot yield infectious, transforming progeny, however, because the sarcoma virus genome cannot specify its own reverse transcriptase.

These experiments appear, therefore, to prove that reverse transcriptase is essential for establishing transformation, just as the provirus hypothesis of Temin predicts, and they also indicate that once a cell has been transformed by a provirus reverse transcriptase is not essential for the maintenance of transformation. Of course, if transformation depends on the integration of provirus DNA, made by reverse transcriptase, into the host genome several other enzymes must be involved in the process. The two nucleases and the DNA ligase which Temin's group have detected in tumour viruses, in addition to the transcriptase (*Nature New Biology*, **230**, 232; 1971), may be there to accomplish the recombination of provirus and host DNA.

While Hanafusa's experiments indicate that reverse transcriptase is required only during the first step of transformation, Martin (*Nature*, **227**, 1021; 1970) has reported the isolation of a mutant strain of Rous sarcoma virus which carries a lesion in a viral gene, the product of which is required continuously to keep the cell in a fully transformed state. Martin exposed Rous sarcoma virus to nitrosoguanidine and then screened the mutagenized stocks for viruses which transform chick fibroblasts at 35° C but fail to transform cells at 41° C. One of the mutants he has isolated has these properties: cells transformed by the virus at 35° C have all the properties of typical transformed cells; when, however, the temperature of the cultures is raised to 41° C the cells rapidly assume many of the characteristics of untransformed cells growing in culture. If the temperature is lowered again to 35° C the cells once more appear to be transformed. Clearly this mutant carries a temperature sensitive lesion in a viral gene which must be continuously expressed to maintain transformation *in vitro*.

Whether or not Martin can show that this mutation also affects the growth of sarcomas induced in chicks in the same way in which it affects the transformation of cells in culture remains to be seen. If he can prove that the mutant affects tumours *in vivo* the identification of the product the mutated gene specifies should answer many of the outstanding questions of tumour biology.

Oncogenes

Although RNA tumour viruses and some DNA viruses (see below) are able to infect and transform susceptible cells in culture and induce tumours in susceptible animals, very few natural cancers have the epidemiological characteristics of infectious diseases. For this reason Huebner and Todaro postulated that genes capable of causing malignant transformation, so called oncogenes, are part and parcel of every animal's genetic make-up. Carcinogens, they envisage, act by switching on oncogenes or in the case of tumour viruses by doing this and also introducing more copies of oncogenes into a cell. Such an all embracing hypothesis is hard to test rigorously but Hanafusa *et al.* (*Proc. US Nat. Acad. Sci.*, **67**, 1797; 1970), Vogt and Früs (Second Lepetit Colloquium, 1970) and Weiss during the past year have provided perhaps the most convincing evidence so far to support the oncogene story. In short, they have been able to isolate an avian leukaemia-type virus (it has yet to be proven that this virus can induce leukaemia) from the cells of apparently healthy chick embryos and Weiss has even induced the production of this virus in the cells of wild Malayan jungle fowl. These observations suggest that the cells of all chickens, irrespective of whether or not the birds are destined to develop leukaemia, have the genetic potential to specify an avian leukaemia virus.

Inevitably, RNA tumour virologists in their first flush of enthusiasm after the discovery of reverse transcriptase have stolen most of the limelight from their colleagues working with the DNA tumour viruses, which outside the laboratory are unlikely to be important carcinogens. But it would be mischievous to imply that the DNA tumour virologists have been gathering moss. Both Eckhart and Dulbecco (*Proc. US Nat. Acad. Sci.*, **67**, 1775; 1970) and Benjamin (*Proc. US Nat. Acad. Sci.*, **67**, 394; 1970), for example, have characterized conditional lethal and host range mutants of polyoma virus which carry lesions in a gene or genes, the product(s) of which are required continuously to maintain a cell in the transformed state. These mutants are equivalent to that of Rous sarcoma virus isolated by Martin (see earlier) and it is now clear that the integration of a tumour virus genome into the genome of a host cell is not sufficient although it appears to be necessary to maintain transformation. And the recent experiments of Wall and Darnell (*Nature New Biology*, **232**, 73; 1971) dispel any remaining doubts about the integration of polyoma virus DNA into host cell DNA. For they have shown that in polyoma transformants messenger RNA molecules are made which contain both cellular and viral base sequences in the same molecule.

But perhaps the greatest stir among DNA tumour virologists during the past twelve months has stemmed from experiments with agglutinins or lectins, glycoproteins which are obtained from plant seeds and which cause cells to clump or agglutinate. It now seems clear that although both normal cells and cells transformed by DNA tumour viruses and some chemical carcinogens bind more or less the same amount of any given agglutinin (Ozanne and Sambrook, Livingstone and Cline, *Nature New Biology*, **232**, 155, 156; 1971) the transformed cells are more readily agglutinated than normal cells. Moreover, cells transformed by the temperature sensitive mutant Eckhart has characterized behave like normal cells when grown at the non-permissive temperature. These findings indicate that the surface of a transformed cell differs from that of its untransformed counterpart and that this change is probably controlled by the transforming gene of the virus. Changes in the cell surface on transformation no doubt account for many of the characteristic properties of transformed cells including their failure to respond to contact inhibition of growth when maintained in culture. And *in vivo* the ability of cancer cells to metastasize no doubt stems from some change in the cell surface. Clearly much future research in tumour virology will be concentrated on the chemistry and structure of the surface of untransformed and transformed cells as well as on reverse transcriptase, the genetics of tumour viruses and the search for human tumour viruses.

Glaciation and the Stones of Stonehenge

G. A. KELLAWAY

Institute of Geological Sciences, 5 Princes Gate, London SW7

More geological evidence can now be brought to bear on the problem of the origin of Stonehenge. Spreading ice sheets rather than human activity could have brought the rocks to Salisbury Plain.

FOLLOWING the discovery of glacial deposits in north Somerset¹, the search for other evidence of regional glaciation has continued. In 1969, shortly before the M5 motorway sections at Clevedon became available, Dr A. Whittaker of the Institute of Geological Sciences found abundant fragments of white Chalk scattered on the steep landslipped slopes of Brent Knoll, Somerset. The Knoll is an isolated hill composed chiefly of Jurassic clay rising out of the coastal marshes between the Mendips and the mouth of the River Parrett and attaining an altitude of 457 feet (140 m) above O.D. This was the first intimation that Cretaceous Chalk might exist in the Drift deposits of Somerset. The source of the Chalk fragments is still uncertain. However, in the light of subsequent developments it seems that they may have been derived from submarine outcrops, or from outliers of Chalk (now perhaps destroyed), the material having been deposited by an ice sheet entering Somerset from the south or west.

Late in the summer of 1970, I found Chalk-bearing Drift deposits at Holwell near Nunney [ST 72554523] in the eastern Mendips where temporary exposures were seen while ground was being cleared for a new quarry. Here reddish brown sandy clay up to 3–4 feet (1 m) thick rests on Jurassic limestone (Inferior Oolite). The deposit is thin and patchy and has been partly decalcified, but the deeper pockets contain rounded and subangular fragments of hard white Chalk thought by Mr C. Wood to have come from the lower or middle part of the formation. Flint, chert, sandstone, quartzite and Carboniferous and Jurassic limestone pebbles are present, suggesting that the material is a relict till.

These and other glacial deposits situated on limestone uplands at the eastern margin of the Mendips at an altitude of about 500 feet (152 m) above O.D., lie in the proximity of gorge-like melt-water channels transecting the Mendip ridge at Nunney and East Cranmore. Their discovery at once suggests that ice could have blocked the western approaches to the Vale of Pewsey and the low lying ground described by F. W.

Harmer² as "Lake Trowbridge". He believed that water from this lake was responsible for the downcutting of the Avon Gorge at Clifton.

Spread of Ice

During the most recent glaciation (Devensian, or Weichsel) ice from the uplands of South Wales reached the Vale of Glamorgan locally, but the main mass of ice in the Irish Sea did not cross the Pembrokeshire peninsula. It may, however, have extended somewhat further south in St George's Channel than was formerly supposed³. Most of the lowlands of South Wales and the greater part of the west and south of England, with the possible exception of Dartmoor and some other high ground, were unglaciated. Winter snowfall may have been heavy in the south-west, permafrost more strongly developed in the south-east of England. The Devensian ice of the Midlands seems to have come no further south than Wolverhampton, thus the greater part of the lower Severn Valley and the Cotswolds were not glaciated. Thick deposits of Head formed in Devon and Cornwall, the Mendips and Cotswolds, and were represented in Wessex and south-eastern England by the familiar "Coombe Rock".

Conditions during the two glaciations preceding the Devensian were very different. The penultimate (Wolstonian, or Saale) glaciation seems to have affected most, if not all, of South Wales⁴ and in the deep valleys of the Pennant Sandstone uplands it is difficult to make a distinction between Devensian and Wolstonian glacial features and deposits⁵. In Pembrokeshire, the Vale of Glamorgan and Somerset the effects of the Wolstonian Glaciation are however recognizable.

Using information accumulated during the past decade the position of the main streams of ice moving eastwards along the Bristol Channel during the Wolstonian Glaciation can now be outlined. In the far west Lundy Island is said to have been overridden by the Wolstonian ice sheet⁶, and glacial deposits were laid down at Fremington and Barnstaple⁷. Further east the great flow of ice which passed over Pembrokeshire was constricted as it moved up channel, confined on the north by the high ground of the South Wales Coalfield⁸ and on the south by Exmoor and the Brendon Hills. Crossing the Severn Estuary, one arm of the ice moved up the broad Flax Bourton valley leading to Bristol¹. Another stream, after dividing round the western Mendips, moved eastwards, crossed the lowlands of Somerset and invaded West Wiltshire. The northern part of the ice flow carried rocks which were derived principally from the South Wales Coalfield and the

Severn Estuary. In the central region erratics which can be matched with rocks from the Lake District, North and South Wales are found. In the south the erratics come mostly from Devon and West Somerset. Considerable quantities of marine and estuarine material torn or scooped up from the exposed floor of the Bristol Channel were carried into Central Somerset and the Clevedon area. Cretaceous flint and chert is abundant in almost all the glacial deposits, and has found its way into the younger raised beach deposits and the Devensian Head. Unlike the much older "Anglian" or Elster Glaciation the Wolstonian ice appears to have introduced no sarsen stones, and where this glaciation was strongest the Tertiary sarsen erratics carried by the older "Anglian" ice have been cleared away or thrust to one side.

"Anglian" Glaciation

The question remains, however, whether the glaciation, which seems to predate the so called "25-foot" raised beach of the Weston-super-Mare district and may therefore be comparable in age with the Chalky Boulder Clay glaciation of the Vale of Moreton, was responsible for the introduction of erratics to Salisbury Plain. The Wolstonian ice would have been capable of conveying Pembrokeshire material to the region east of the Severn Estuary and Bristol Channel. There are, however, indications of an older "Anglian" glaciation of the West of England and Bristol Channel which was much more extensive and carried very large Scottish erratics weighing as much as 50 tons⁸. This glaciation was probably responsible for the introduction of Welsh erratics to Salisbury Plain (Fig. 2). It equates with the ancient Plateau Gravel or "Northern Drift" ice sheet which carried Bunter pebbles and other northern erratics into the Cotswolds and the Upper Thames Valley and which preceded the formation of the Hanborough Terrace of the Upper Thames and Evenlode Valleys⁹. The magnitude of the glaciation can be judged from the fact that the upper surface of the ice-sheet in the vicinity of Cheltenham stood at well over 1,000 feet (305 m) above present sea level. One of the first comprehensive studies of the problems presented by the fragmentary record of this great glaciation, of which little evidence has survived further north, was made by Sandford^{10,11}.

The "Anglian" glaciation of the eastern part of the English Channel is thought to be of the same age as the "Anglian" glaciation of Devon and Cornwall where the presence of big "foreign" erratics has been attributed to the action of "coastal ice". The situation of the big erratics found at Croyde in North Devon cannot be explained in this way without assuming that the sea level was as high or higher than at present; this during a major glaciation. As Taylor has pointed out^{12,13}, the presence on the north coast of Devon of rocks from South Devon can hardly be accounted for by the drifting ice theory without arousing incredulity.

The name "Anglian" is in many ways unsuitable because of its link with the East Anglian succession. It should therefore be clearly understood that, as used in this account, the term refers to the ice sheet which invaded the coastal region of Hampshire and Sussex from the west carrying sarsens and big "foreign" erratics^{14,21}. Erratic pebbles from these or comparable glacial deposits were later incorporated in the marine gravels of the Slindon (or 100-foot) raised beach of West Sussex¹⁵. Others found their way into the younger solifluction and beach deposits of the coastal plain.

The very low sea level accompanying the "Anglian" glaciation of southern England and the English Channel was therefore separated by the high Slindon (Tyrrhenian) sea level from that of the Wolstonian (Saale) glaciation. During the period of falling sea level marking the retreat of the shore-line from the 100–300 foot level intensive piping and disturbance of the Slindon raised beach and the formation of very thick solifluction deposits (the Main Coombe Rock of Sussex) took place.

It is generally considered, however, that the ice of the

Wolstonian glaciation did not extend into the central and eastern parts of southern England. Salisbury Plain, though dissected by melt-water from the Wolstonian ice, does not seem to have been covered by an ice sheet at this time. The removal of the Clay-with-flints which was noted by Reid¹⁶, probably took place during the "Anglian" glaciation. The detailed movements of the various ice masses invading southern England during this glaciation are as yet unknown. They may well have been complex, and the directions shown in Figs. 1 and 2 should be regarded as only an approximation.

It seems likely that the older highly cryoturbated Plateau Gravels of the Hampshire Basin represent disturbed glacial outwash from the "Anglian" ice since they are composed of material derived from the Clay-with-flints, and the local Mesozoic and Tertiary rocks. Small and medium-sized sarsen boulders are present, and Palaeozoic rocks also occur particularly in the western districts.

Irish Sea Erratics

Turning to the question of the known eastern limit of Irish Sea erratics, pebbles of porphyritic igneous rocks together with flint, chert and "flinty slate" were found in the nineteenth century on Flat Holm island in the Bristol Channel¹⁷. The Geological Survey collection includes a fragment of a "soda-granite" from the same locality, presumably part of a pebble of Welsh origin. The desirability of re-examining Flat and Steep Holm has been evident for some time and in May 1971 I paid a brief visit to the islands in company with Mr Brian Hawkins of Bristol University and Dr Brian Kelk of the Institute of Geological Sciences.

On the south side of Flat Holm the storm beaches are composed almost entirely of local Carboniferous Limestone (Goblin Combe Oolite and Clifton Down Mudstone) but on the north side the limestone pebbles are mixed with Chalk flint, greensand chert, Old Red Sandstone and a wide range of igneous and metamorphic rocks which have been identified by Mr R. K. Harrison and seem to have been derived from South Wales, North Wales and the Lake District. The erratic tracks fit with the evidence from South Wales where Bowen⁴ has recently summarized the work of Griffiths¹⁸ demonstrating the dispersal of Pembrokeshire and other rocks carried by ice from the Irish Sea, probably in Wolstonian times. On the other hand, the Flat Holm beach material includes very little rock from the mainland coasts and the most likely explanation is that the pebbles are derived from a submerged deposit of till, probably lying off the north coast of the island. Further discussion is postponed until this has been located and other reported occurrences of erratics have been investigated.

Here it may be useful to observe that reddish brown till with Chalk flints, greensand chert and rocks from South Wales was found in 1970 at or below sea level at Kenn near Clevedon, though in this case the glacial deposits are buried by Holocene mud¹. Other residual masses of glacial sand and gravel ("Burtle Beds") rise out of the marshes in central Somerset at Catcott Burtle, Westonzoyle and Othry¹⁹. In these outwash deposits, previously described as marine, the remains of marine, estuarine and freshwater organisms as well as the bones of land animals occur. The associated erratics are derived chiefly from West Somerset and North Devon. They include Devonian or Old Red Sandstone, Carboniferous, Bunter and Jurassic rocks, Cretaceous chert and flint. These are associated with pellets of soft red and green Triassic mudstone, some mixed with marine shells and remoulded in ice, others, lumps which may have been rounded while in a frozen condition. At least one mega-erratic is thought to have existed in the Westonzoyle area²⁰. This was known as the "Devil's Upping Stock". Unfortunately it was cut up for ornamental stone and its petrological affinities are therefore unknown.

Scattered on the Dorset, Marlborough and Berkshire downs are isolated groups of Tertiary sandstone erratics known as sarsens. The name is probably derived from the Anglo-Saxon

sar (troublesome) and stan (a stone), and, as Chatwin²¹ has pointed out, it is significant that a farming people should have paid particular attention to them.

Many differing views have been expressed as to their ultimate source and mode of emplacement. A full discussion of these problems must await more detailed description. For present purposes, however, it may be said that in certain areas, notably Sussex and Somerset, they behave as would any other erratics and are associated locally with erratics of "foreign" rocks which seem to have travelled with them. The Wolstonian ice of the Bristol Channel does not seem to have picked up Tertiary sandstone, possibly because it did not traverse the extensive tracts of submerged Tertiary rocks found in the English Channel and the western approaches. Nor were

chalk, progress in understanding the distribution of the sarsens and the glaciation of southern England was impeded.

One of the most mysterious features of the great megalithic monument of Stonehenge is the presence of the foreign stones, or "bluestones" as they are popularly described²². It has long been known that these are not of local origin and there has been intense speculation about their presence on Salisbury Plain. Since the bluestones of Stonehenge are considered to have come from Pembrokeshire, the discovery of evidence for the glaciation of the region east of the Bristol Channel is clearly of more than local geological significance. H. H. Thomas²³ who was the first to recognize in Pembrokeshire the probable source of the bluestones, believed that there was no valid evidence to support the theory advanced by Judd²⁴,

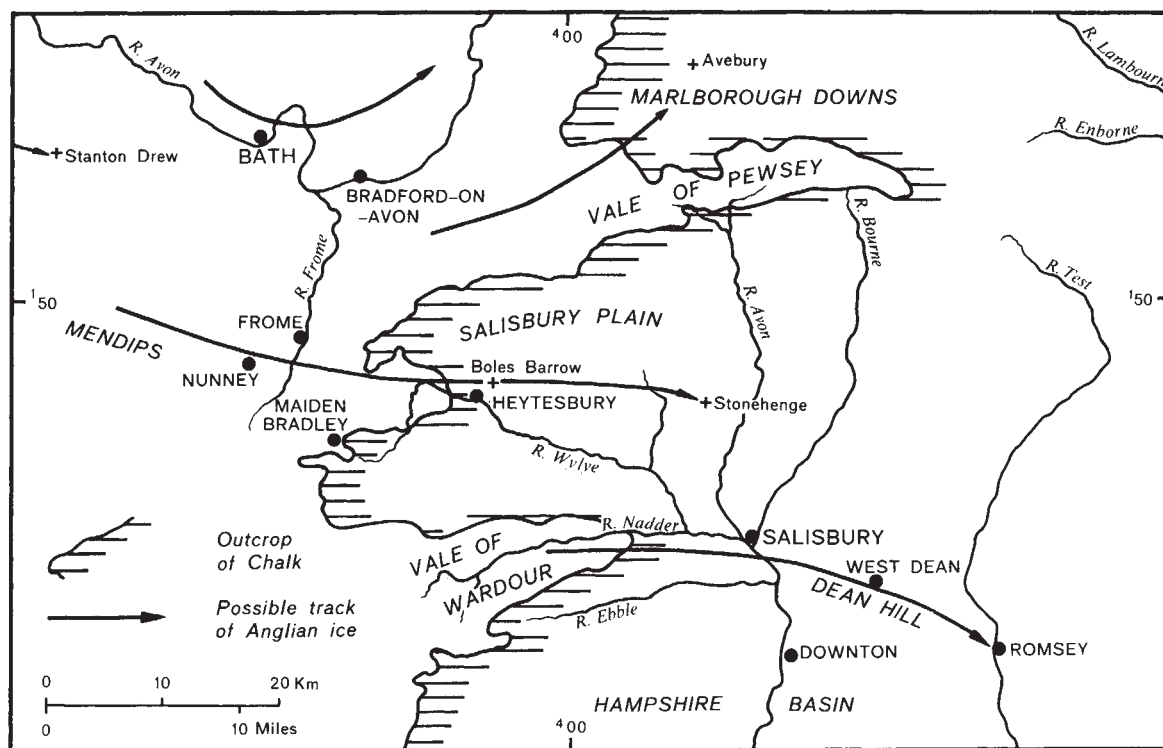


Fig. 1 Sketch map of Salisbury Plain and its environs showing principal localities referred to in the text. Ice entering Wiltshire from the west may subsequently have been deflected by sarsen-bearing ice from the English Channel moving in a SW-NE direction.

sarsens carried by the "Northern Drift" ice of the Cotswolds and Severn Valley.

South of Bristol and Bath sarsen erratics are found on high ground in the eastern Mendips and Broadfield Down, mostly in areas which were not eroded by the Wolstonian ice. In Sussex, on the other hand, the sarsen erratics were obviously brought in by ice from the English Channel and the most likely sources of the sandstone, which is not found in the local rocks on shore, are the Tertiary formations which are known to cover large areas of the sea floor.

Sarsens and Stonehenge

For the archaeologist the sarsen stones and their origin are matters of particular interest, since the two greatest megalithic monuments in the British Isles, Avebury and Stonehenge, are built wholly or in part of sarsen sandstone. As long as it was believed that the sarsens were the remains of a relict Tertiary sandstone "lowered" into position by the dissolution of the

that they were conveyed to Salisbury Plain by ice. At this time the existence of glacial deposits at Fremington in North Devon had only tentatively been recognized by Maw (1864)²⁵ and their glacial origin was not generally acknowledged until much later^{18,26,27}. On the other hand, the clearly expressed views of Harmer² respecting the glacial origin of the gorges of the Bristol Avon were widely ignored or frankly disbelieved by most geologists working in the area. Thomas therefore concluded that ice from the Irish Sea did not reach Devon and Somerset and went on to say that the available evidence "permanently disposes of the idea of glacial transport for the Foreign Stones of Stonehenge".

The effect of this pronouncement on archaeology and, indirectly, on English glacial geology has been considerable. H. H. Thomas explained the presence of the foreign stones on Salisbury Plain and the wide variety of rocks represented among the bluestones by suggesting that they were originally erratics gathered up locally and used in the construction of a great stone circle which he thought might have been built at

Cil-maen-llwyd on the south-eastern slopes of the Preseli mountains²³. In this way he accounted for the presence of relatively soft sheared volcanic rocks which would not otherwise have been used for building, particularly where hard dolerite and rhyolite are freely available as they are at Mynydd Preseli. According to this theory the stones were subsequently removed from this "venerated stone circle" and conveyed by its votaries to Stonehenge. The archaeological evidence for the "venerated stone circle" has since been discredited^{28,22}; but belief in the human transport of the stones has persisted, though the theory of overland transport preferred by H. H. Thomas has been replaced by carriage by water.

One outstanding item in this problem is the Altar Stone, shown by Thomas to have been derived originally from the

to exist. Both are preserved in Salisbury Museum. One is a boulder weighing about 9 ton found in company with sarsen stones in the Neolithic long barrow known as Boles Barrow near Heytesbury [ST 94214677]. The other is a smaller stone from Lake, south of Amesbury in the valley of the Salisbury Avon. Both appear to be petrologically similar to the dolerites from Mynydd Preseli and those at Stonehenge. The stone from Lake may, however, be a piece detached from one of the Stonehenge bluestones in comparatively recent times and for present purposes is best ignored.

If the bluestones of Stonehenge were conveyed in ice from Pembrokeshire it might reasonably be expected that some ice-scratched or polished erratics would have been found in Wiltshire. Many of the bluestones used at Stonehenge have

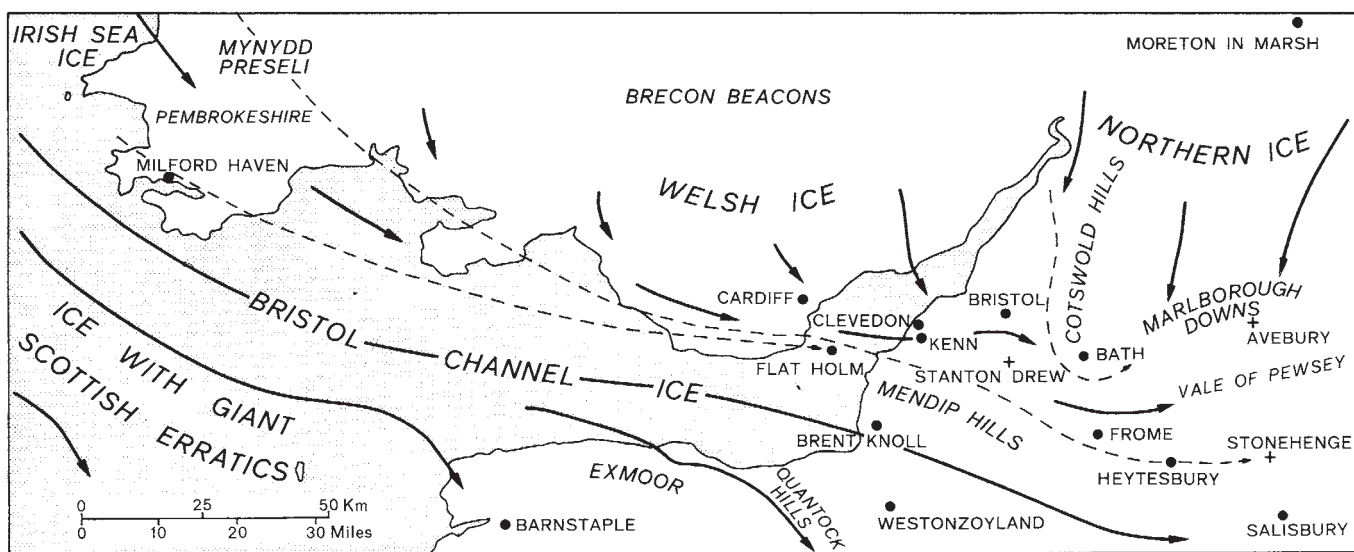


Fig. 2 Bristol Channel region showing generalized erratic "tracks" and inferred direction of flow of the "Anglian" (Elster) ice. Exmoor may have formed a small centre of local ice accumulation, the Mendip and Cotswold Hills were largely overridden at the height of the glaciation. The northern limit of the sarsen erratics is given roughly by a line joining the southern end of the Quantock Hills, Stanton Drew, Bath and the northern edge of the Marlborough Downs.

Senni Beds of Breconshire or, alternatively, from the Cosheston Sandstone of southern Pembrokeshire. Even if we suppose that all the Old Red Sandstone material found at Stonehenge came from South Pembrokeshire, there is still no means of establishing Mill Bay or Milford Haven as the immediate source of the material used at Stonehenge. Not only does the Old Red Sandstone of south Pembrokeshire cover a fairly large area on land, it also extends beneath the sea and it has long been known that the whole region has been strongly glaciated. It may be that Neolithic man picked up a 7 ton boulder of Cosheston Sandstone on the foreshore of Milford Haven. On the other hand, ice could have quarried the bedrock or picked up a boulder from a very much wider area, and deposited it anywhere between south Pembrokeshire and Salisbury Plain. It is certain that neither the Altar Stone nor some of the other sedimentary rocks recognized by Judd¹⁷ at Stonehenge came from the Preseli Mountains. As Judd pointed out, many of the Palaeozoic rocks found at Stonehenge would have been too soft or fissile to have been of any value as a building material. The inference is plain. Even if the hard igneous rocks were worth transporting to Stonehenge, the softer stones have still to be accounted for, and glacial action is the only reasonable solution.

Apart from the bluestones at Stonehenge, only two other large pieces of igneous rock from Salisbury Plain are known

been dressed or shaped by hand, others show rough or matt surfaces which convey no useful information beyond the fact that they have undergone some weathering. The Boles Barrow stone does appear to retain part of its original surface on which the harder crystalline knots or patches have been faceted. This may be a relic of glacial smoothing, but much of the polish on the stone is due to handling.

The relative age of Boles Barrow and Stonehenge is said to be uncertain, though the former is attributed to the Windmill Hill culture and Stonehenge to the Beaker culture. Archaeological evidence shows that the foreign stones were certainly on Salisbury Plain in Neolithic times²⁹ and that the bluestones of Stonehenge were originally used in the construction of an earlier monument, possibly at the Cursus³⁰.

Other Means of Transport

Following rejection of the theory of glacial transport, carriage of the stones on boats or rafts by way of the strongly tidal Severn Estuary and Bristol Avon has come to be the most widely accepted hypothesis. Nevertheless the Neolithic and Bronze age rivers differed considerably from those of the present day, and the Bristol Avon has been virtually converted into a man-made channel by centuries of excavation, dredging and control by embankments, lock gates and weirs. It seems

to me questionable whether river transport of such massive stones was practicable above Bathampton in Neolithic times even if the serious navigational hazards of the Severn Estuary and the lower Avon had been surmounted.

Why should the builders of Stonehenge carry lumps of slate, quartzite, greywacke, flagstone, soft sheared volcanic ash and Old Red Sandstone to Stonehenge when ample supplies of hard dolerite were available at Mynydd Preseli? The strange mixture of stones—Old Red Sandstone, dolerite, rhyolite, volcanic ash, calcareous ash—described by Atkinson²² from the west side of the bluestone circle, and the presence of Chilmark Stone from the Vale of Wardour in the packing of stone holes, certainly supports Judd and Thomas in their conclusion that the stones were originally erratics. The reasons for thinking that these were collected from a “venerated stone circle” and conveyed overland are, as we have seen, no longer acceptable. The suggestion that Mynydd Preseli was a “sacred mountain” leaves unexplained the use of Cosheston Sandstone and other Old Red Sandstone rocks from unsanctified south Pembrokeshire. Yet, if the stones did not come from one “venerated stone circle” or “sacred mountain” the motives for conveying them 150 miles (241 km) to Stonehenge while ignoring local resources, are incomprehensible. For large masses of suitable stone (Great Oolite Limestone) remain to this day, within 20 miles (32 km) of Stonehenge on the land-slipped slopes of the Avon valley between Bathampton and Bradford-on-Avon, the identical defile through which the Pembrokeshire rocks are supposed to have been conveyed by Bronze Age man to Salisbury Plain.

In view of recent developments a re-examination of the Stonehenge rocks has been commenced. Whatever may be the outcome of these investigations there is bound to be an element of doubt about some of the earlier conclusions. While it is true that some or all of the igneous bluestones could well have come from Mynydd Preseli, comparable rocks are known in the Ordovician terrain near Cader Idris^{23,31}. There may also be offshore areas of similar Ordovician volcanic and associated intrusive rocks in the Irish Sea and St George's Channel¹⁸.

This does not affect the interpretation of the glacial history since it can be demonstrated that ice flowed out of the Irish Sea into the Bristol Channel carrying rocks from a very wide region. However, it has a critical bearing on the theory of human transport which depends on the material having been derived from a limited area on land. If it could be proved that Neolithic man actually quarried the crop of the spotted dolerite at Carn Meini or worked the frost-shattered debris on Mynydd Preseli the theory of human transport might be credible. No such evidence has been produced. Neolithic flint miners dug shafts and galleries in the chalk in Sussex and East Anglia but their tools, like those used at Stonehenge, were deer horn, bone, flint, sarsen and wood, of little value for quarrying hard massive rocks. In practice, Neolithic man therefore used lumps and boulders, or tabular erratics and slabs that he could split with the aid of fire and water.

Frost-wedged fragments, though not strictly erratics, could also have been used at Stonehenge, but even if this could be established it would not prove transport by human agency, since some of the Stonehenge erratics may well have been broken down *in situ* by frost action on Salisbury Plain.

Archaeological attention has been diverted from the effects of glaciation and concentrated on the petrology of the igneous bluestones in the belief that it is possible in this way to prove that they were conveyed by man directly to Stonehenge. Without questioning the conclusions of Thomas²³ concerning the original source of the igneous bluestones used at Stonehenge, it must be admitted that there are circumstances when the use of specific “indicator rocks” to pinpoint individual localities can be misleading. Of very few rocks can it be said that they are unique and confined to one small area and that no other occurrences may exist on the sea floor or beneath drift deposits on land. Pebbles and fragments may be secondarily derived,

glacial erratics can be introduced by an early glaciation and moved again during a later one. Errors in determining provenance are one of the hazards of using erratics but the risks have to be accepted if progress is not to cease

Erratics in Megalithic Monuments

If Stonehenge and Boles Barrow are constructed of, or incorporated, large erratics deposited by the “Anglian” ice, other megalithic monuments in the area glaciated by the “Anglian” and Wolstonian ice may be expected to show similar features. For this reason the stone circles of Stanton Drew and the outlying stones in the Chew valley of north Somerset³² are an ideal case for investigation. The monument is sited on Triassic marl but local (Triassic) sandstone is present in the vicinity and would have been suitable for quarrying if Neolithic or Bronze age man had been able to do so. Undressed stones up to about 27 tons in weight, but of variable size and shape, have been used at Stanton Drew, none of them of local origin. The monument includes three stone circles, and a group known as The Cove. Some of the outlying stones, including Hautville's Quoit, may or may not have formed part of a structure. Rocks used in the monument or found at the Cove and Hautville's Quoit include silicified and mineralized dolomitic conglomerate from Broadfield Down and the Mendips³³, Old Red Sandstone possibly from Wales or from the Portishead Beds of Blackdown or the Portbury district, Dundry Freestone (Inferior Oolite) and Tertiary sarsen stone. The massive blocks of Dundry Freestone can only have come from the western end of Dundry Hill. The sarsen, which it has not yet been possible to examine petrologically, can be matched by sarsen fragments found on Felton Common [ST 518651] the broad west-east valley crossing Broadfield Down 6.5 miles (10 km) south-west of Bristol. Fragments of sarsen stone at Felton Common were previously mistaken for Rhaetic (which also occurs on Broadfield Down), but the rock contains the characteristic rootholes. Sarsen erratics, presumably introduced by the “Anglian” ice, are not uncommon on high ground south of Bristol. These may not have been over-ridden by the Wolstonian ice sheet.

Tracks of the stones used at Stanton Drew have a strong west-east component and it is noticeable that though Triassic sandstones are present locally and hard Coal Measures sandstones crop out about 1.5 miles (2.4 km) east of Stanton Drew, neither of these rocks has been used in the construction of the megalithic monument. The situation of the stones of Stanton Drew is entirely consistent with the possible site of a moraine upstream of the Chew gorge at Pensford [ST 617637]. The latter may have operated for a time as a spillway from a glacial lake covering the upper part of the Chew valley, including the area of the present reservoir. In this instance the erratics were probably introduced by the “Anglian” ice, final emplacement having taken place in Wolstonian times. Some of these large tabular erratics, notably the dolomitic conglomerate of The Cove, could have been in the form of giant slabs which were later broken up, either by natural processes or by man.

Other large isolated stones or groups of stones formerly existing in otherwise stoneless areas of Wessex have been broken up for hard core or destroyed as pagan monuments. Some of these acts of vandalism have been inscribed but the only record of the existence of other vanished stones is to be found in the local place names or in folklore³⁴. The surviving greywethers in the downlands of Wessex and Sussex, the Mendips and west of England, probably give only a faint impression of the former extent of the boulder strewn surfaces.

Many erratics are almost certainly enclosed in chambered long barrows and similar structures and a comprehensive study of these rocks might yield valuable evidence bearing on the direction of “Anglian” ice movements. Grinsell³² has estimated that there are about 200 long barrows in Wessex alone. Not all of them are chambered, and the average weight of stone per barrow can only be guessed, but if 100 long barrows

contain an average of only 50 tons a piece it would mean that 5,000 tons of medium-large erratics could be included within them. This would still leave uncounted the enormous quantities of stone originally used in sites such as Avebury where some of the sarsen erratics weigh up to 40 tons and one of 90 tons formerly existed.

Erratic boulders of up to 350 tons are said to exist in the Carnac area of south Brittany and giant erratics and glacial rafts of huge dimensions are found in Britain and many parts of the world. Nobody knows the original size of the Stonehenge sarsen erratics before they were split and trimmed, but their stratigraphy suggests that a number of them may have formed parts of one or more giant slabs stranded somewhere in the vicinity of the monument. As an example demonstrating the size which can be attained by such glacial erratics we may refer to the immense mass of tabular Silurian limestone found in Warren County, Ohio. This slab, though only 5 feet (1.5 m) thick has a surface area of 20,000 square feet (1,858 m²)³⁵. Even larger erratics consisting solely of quartzite are found in a boulder train near Calgary, Alberta³⁶. Beneath such huge tabular masses, till, pebbles and boulders may be preserved even when the surrounding areas have been cleared by erosion. This may well be the explanation of the survival at Stonehenge of lumps of relatively soft volcanic rock such as the ones represented only by stumps in the bluestone circle. Soft rocks, or compact but frost-susceptible erratics partly buried in glacially disturbed Chalk or chalky till and preserved from erosion beneath huge flat-lying frost-resistant sarsen stones, may have been unearthed by Neolithic or Bronze age man when building his monuments. Some of these smaller erratics may have appeared to be hard enough to use for building and may have survived for a time before breaking down under the influence of frost and rain. Without some such explanation it is difficult to see why the rocks should ever have been used at all, much less carted from Pembrokeshire and then trimmed on site with all the attendant risk of wasting stone.

It has also been claimed that the sarsens of Stonehenge were dragged overland from Avebury, but there is no geological evidence to support this. Nor does it seem at all likely that they were moved on frozen ground in late Neolithic or early Bronze age times³⁷, when the climate may have been warmer than at the present day.

There seems to be a close correlation between the distribution of long barrows in Wessex (ref. 32, Map 1) and the downland areas now thought to have been invaded by the vast but ancient "Anglian" ice sheet. For example, the presence of Jurassic limestone built into the chambered West Kennet Long Barrow [SU 10506775] and other barrows in the Bishop's Cannings valley on Marlborough Downs^{32,38} suggests that the Avebury sarsens may have been dumped as part of a moraine at the terminus of a stream of ice moving in an easterly or north-easterly direction (Fig. 1). It is not unreasonable to suppose that the presence of non-local clay found in the centre of some barrows may be the remains of till deposits disturbed during the lifting and erection of erratics.

Much work will be needed on the chambered long barrows before detailed conclusions can be drawn about these movements, and in many areas there is little reliable information as to the nature of the stones used in their construction. It may be no coincidence, however, that the western part of the Marlborough Downs has some of the most characteristically glaciated profiles to be seen in this part of Britain. The moulded outlines, broad U-shaped dry valleys and melt water channels of the downlands are in some respects the most interesting of the glacial features, but their assessment demands more extended treatment than is possible at this stage.

Any calculation of the original quantity of erratic material on Salisbury Plain necessarily ignores the smaller pebbles and fine grained material dissolved, blown, washed away, or moved by solifluction during dissection of the uplands in the two major cold periods following the "Anglian" glaciation. To

these later periods belong the valley gravels of the Wylye, Avon, Test and other rivers, and the solifluction streams which concentrated the medium and some of the big sarsens in the dry valleys of the downlands. Only the inertia of the larger erratics and their situation on normally waterless uplands preserved them for the use of Neolithic man.

I thank everybody who has provided information and material, especially Mr Brian Hawkins, Dr Brian Kelk, Mr R. K. Harrison and Mr C. Wood. I also thank Mr H. S. Shortt, Mr C. Bowen and Mr R. S. Newall for help with archaeological problems. The manuscript was edited by Miss E. M. Pyatt.

- ¹ Hawkins, A. B., and Kellaway, G. A., *Proc. Geol. Assoc.*, **82**, No. 2, 267 (1971).
- ² Harmer, F. W., *Quart. J. Geol. Soc.*, London, **63**, 470 (1907).
- ³ John, B. C., *Geol. Mag.*, **107**, 439 (1970).
- ⁴ Bowen, D. Q., in *The Glaciation of Wales and Adjoining Regions* (edit. by Lewis, C. A., London and Colchester, 1970).
- ⁵ Woodland, A. W., and Evans, W. B., *Geology of the South Wales Coalfield, IV, The Country around Pontypridd and Maesteg.*, *Mem. Geol. Surv.* (1964).
- ⁶ Mitchell, F., *Proc. Roy. Geol. Soc., Cornwall*, **20**, No. 1 for 1956-6, 65 (1968).
- ⁷ Edmonds, E. A., *Rep. Inst. Geol. Sci.* (in the press).
- ⁸ Flett, J. S., *The Geology of the Lizard and Meneage*, *Mem. Geol. Surv.* (1946).
- ⁹ Kellaway, G. A., Horton, A., and Poole, E. G., *The Development of some Pleistocene Structure in the Cotswolds and Upper Thames Basin*, *Bull. Geol. Surv.* (in the press).
- ¹⁰ Sandford, K. S., in Pringle, J., *The Geology of the Country around Oxford*, *Mem. Geol. Surv.* (1926).
- ¹¹ Sandford, K. S., *Quart. J. Geol. Soc., Lond.*, **85**, 359 (1929).
- ¹² Taylor, C. W., *Trans. Devon. Assoc.*, **88**, 52 (1956).
- ¹³ Taylor, C. W., *Trans. Devon. Assoc.*, **90**, 187 (1958).
- ¹⁴ Reid, C., *The Geology of the Country around Bognor*, *Mem. Geol. Surv.* (1897).
- ¹⁵ Calkin, J. B., *Proc. Prehist. Soc. East Anglia*, **7**, No. 3, 333 (1934).
- ¹⁶ Reid, C., *The Geology of the Country around Salisbury*, *Mem. Geol. Surv.* (1903).
- ¹⁷ Buckland, W., and Conybeare, W. D., *Trans. Geol. Soc.*, No. 2, **1**, 210 (1824).
- ¹⁸ Griffiths, J. C., *The Glacial Deposits West of the Taff*, thesis, Univ. London (1940).
- ¹⁹ Bulleid, A., and Jackson, W., *Proc. Somerset Archaeol. Nat. Hist. Soc.* for 1937, **83**, 171 (1938).
- ²⁰ Dobson, D. P., *The Archaeology of Somerset* (Methuen, London, 1931).
- ²¹ Chatwin, C. P., *The Hampshire Basin and Adjoining Areas*, *British Regional Geology*, *Mem. Geol. Surv.* (1948).
- ²² Atkinson, R. J. C., *Stonehenge* (Penguin, London, 1900).
- ²³ Thomas, H. H., *Antiquaries*, **3**, 240 (1923).
- ²⁴ Judd, J. W., in Gowland, W., *Archaeologia*, Second Series, **8**, 106 (1902).
- ²⁵ Maw, G., *Quart. J. Geol. Soc. London*, **20**, 445 (1864).
- ²⁶ Dewey, H., *South-West England*, *British Regional Geology*, *Mem. Geol. Surv.* (1935).
- ²⁷ Zeuner, F. E., *The Pleistocene Period* (Hutchinson, London, 1959).
- ²⁸ Grimes, W. F., *The Prehistory of Wales* (National Museum of Wales, Cardiff, 1951).
- ²⁹ Newall, R. S., *Stonehenge, Wiltshire* (HMSO, London, 1959).
- ³⁰ Stone, J. F. S., *Archaeol. J.*, **104**, 7 (1948).
- ³¹ Cox, A. H., and Wells, A. K., *Quart. J. Geol. Soc. London*, **76**, 254 (1921).
- ³² Grinsell, L. V., *The Archaeology of Wessex* (Methuen, London, 1958).
- ³³ Morgan, C. L., *Proc. Somerset Archaeol. Soc.*, **33**, No. 2, 37 (1887).
- ³⁴ Crawford, O. G. S., *Mounds, Standing Stones and other Objects, The Long Barrows of the Cotswolds* (Gloucester, 1925).
- ³⁵ Flint, R. F., *Glacial Geology and the Pleistocene Epoch* (Wiley, New York and London, 1947).
- ³⁶ Stalker, A. MacS., *The Erratics Train Foothills of Alberta*, *Geol. Surv. Canad. Bull.*, **37** (1956).
- ³⁷ Hawkins, G. S., *Stonehenge Decoded* (London and Glasgow, 1970).
- ³⁸ Piggott, S., *The West Kennet Long Barrow, Excavations 1955-6* *Min. of Works Archaeological Reports*, No. 4 (1962).

Was there really a Big Bang?

G. BURBIDGE

University of California, San Diego

The evidence in favour of a big bang cosmology is much less definite than is widely realized, and it is not impossible that we are living in a steady state universe.

In any given field of science, activity has its peaks and its valleys. Many view activity as synonymous with progress. If one judges by the number of papers currently appearing and by the interest of the younger astrophysicists, activity in cosmology is at a peak at present. To the outsider and even to many inside, it looks, at first sight, as though considerable progress has been made in recent years, so that the outline of the way in which the universe has evolved is understood and only the details need to be explained. Is there any justification for this? Those who have read widely in the earlier and more recent cosmological literature, as well as those who rely on North's account¹, are aware that views of cosmology at any epoch are largely determined by the ideas of a few strong individuals, rather than by an objective appraisal of the information available. Today a similar situation prevails, and it is further complicated because we are often trying to use classes of objects to test cosmological theories long before we understand the physics of them.

In my view, modern developments have been influenced by several factors. They include new observational discoveries combined with an extremely simple minded approach to them, a deep seated conviction by many that general relativity is correct in all details, a belief that astronomy has nothing to teach us about fundamental physics, and, last but not least, a deep seated hostility by both observers and theoreticians in the astronomical community toward the steady state theory. Usually this last point is not mentioned in polite society, but it is blatantly obvious in Dingle's remarks² while president of the Royal Astronomical Society; it is often apparent in private discussions, and a further manifestation is a complete disregard of the theory when the state of cosmology is reviewed (see, for example, ref. 3).

What are the major scientific events which make up the case for a big bang? They started, of course, with the work of Einstein in 1917. Incidentally, he, as did his contemporaries, took it for granted that the universe in the large is unchanging, that is, he assumed that it is in a steady state. But it was Lemaitre^{4,5} and, independently, Friedmann⁶ who showed that the Einstein universe is unstable so that, when perturbed in the direction of expansion, it will continue to expand, tending to a de Sitter universe in the limit. Lemaitre's analysis showed that there is an infinity of models satisfying Einstein's equations. Similar solutions are obtained in the scalar-tensor theory. The most interesting model is the one which starts at a finite time in a state of infinite density. It was therefore inferred that, according to the theory of general relativity, the universe must be expanding, and most likely began in a state of high density. Almost immediately following this major prediction of general rela-

tivity, Hubble and Humason^{7,8} conclusively established the redshift-apparent magnitude relation, which gave observational proof that the universe is expanding. Thus by 1930 Einstein's theory had made what many considered to be the greatest successful prediction in the history of science. For the next twenty years the drive in observational cosmology, almost completely in the hands of Hubble and his colleagues, was to attempt to decide from observation which of the myriad of general relativistic models best represent the universe. Soon, however, there developed an apparent paradox associated with the ages of different systems, because it appeared that the ages of various astronomical objects, notably the Earth, seemed to be greater than the age of the universe derived from the rate of expansion obtained by Hubble. Two possible paths of escape from this dilemma were the Lemaitre model with its arbitrarily long waiting period, and the steady state cosmology requiring the continual creation of matter^{9,10}.

The present situation with regard to this age difficulty is that we now realize that the uncertainties associated with the determination of H and those associated with obtaining the ages of the oldest stars are great enough that we no longer have any strong evidence that there is any problem. Most recently it has been stated that $H \sim 50 \text{ km/s Mpc}^{-1}$, or $H^{-1} \approx 20 \times 10^9$ years^{11,12}, while the age of an old galactic cluster, NGC 188, is given as 10×10^9 years¹³⁻¹⁵, and the age of the elements obtained from the radioactive isotopes is $\sim 10 \times 10^9$ years, with uncertainties depending on assumptions made about the sequence of events leading to the isolation of U and Th in the solar system¹⁶.

The prediction that the universe is expanding, followed by the discovery of the expansion, was the first major success of cosmology based on general relativity. The second came from the comparatively elementary discussion of the physics of the early stage in a dense expanding universe—the primaeval fireball—by Gamow and his colleagues^{17,18}, and the discovery in 1965, by Penzias and Wilson¹⁹, of background radiation which may well be relic radiation from this fireball. These two predictions, followed by two major discoveries which, at first sight, appeared to confirm the theory, provide the basic scientific case for evolving models of the universe.

If we wish to account for the expansion without invoking an initially dense phase, we are left with either the steady state model in which matter is created or with an oscillating model. The steady state alternative seemed attractive to some, especially "when taken in conjunction with aesthetic objections to the creation of the universe in the remote past. For it is against the spirit of scientific enquiry to regard observable effects as arising from causes unknown (and unknowable) to science, and this in principle is what creation in the past implies"¹⁰. There is apparently a great gulf between those who find creation in the past acceptable, but continuous creation unacceptable, and those who have no such strong feelings. It is my belief that much of the conflict in cosmology in modern times has arisen because of this division which has little to do with science. More to the point, we must critically evaluate the detailed evidence which has accrued in the past twenty years concerning the question of whether or not we have definitely established that the universe is evolving.

Determination of q_0

In 1956 Humason, Mayall, and Sandage²⁰ gave a detailed account of the status of the Hubble relation. At that time the largest redshifts which had been measured were $z \approx 0.2$. In the fifteen years that have elapsed since then, Sandage and his collaborators have worked intensively on the problem of refining the measurements of H and q_0 . Many progress reports on this work have been given by Sandage²¹⁻²⁵; see also Peach²⁶. As far as the value of H is concerned there is some reason to believe that we are beginning to converge toward a realistic value some ten times smaller than that obtained by Hubble in 1936. As far as the determination of the deceleration parameter q_0 is concerned, the position is not so good. There are two basic reasons for this. First, very few clusters with redshifts greater than 0.2 have been discovered. The most recent plot given by Sandage²⁵ contains forty-two first-ranked galaxies in clusters, and only three of these, with $z = 0.29$, 0.36 and 0.46, are in excess of $z = 0.2$. The three clusters were observed by Baum²⁷ more than ten years ago, and a detailed discussion of his data has never been given. Many people believe that the radio astronomer's ability to detect more distant objects will lead, or has led, to an improvement of the situation. The only major achievement of radio astronomy in this connexion so far was the identification, in 1960, of 3C 295 with a galaxy with a redshift of $z = 0.46$ (ref. 28). This source lies in a cluster, and Baum obtained photometry which, combined with the redshift, gave this last point on the Hubble diagram.

About 100 galaxies associated with radio sources and 200 quasi-stellar objects, most of which are radio sources, have been optically identified, and redshifts have been obtained. None of these, with the exception of the cluster around 3C 295, has given us any information about the classical Hubble diagram beyond $z \approx 0.2$. This is because the radio sources are associated with galaxies with redshifts ≤ 0.2 , or they are associated with galaxies which obviously cannot be treated as "standard candles", or they have been identified with QSOs which show a very large dispersion in the redshift-apparent magnitude diagram. Using the radio properties as a guide, we can single out a certain class of objects, but it is not clear at all that they can be used for cosmological investigations of the classical kind. If we make a table of optical types of radio galaxies as a function of increasing redshift²⁹, we see that, at small redshifts, almost all types of galaxy are being identified. At intermediate redshifts the galaxies are mostly giant ellipticals, while at the largest redshifts, between 0.2 and 0.3, the galaxies with published redshifts have nearly all been classified as N-type, meaning that a large part of the optical energy that they are emitting is due to non-thermal processes and hot gas rather than stars. Thus they cannot yet be used as standard candles, as the data available at present are inadequate.

Redshifts of Faint Sources

So far it has not been possible to make optical identifications and obtain redshifts of radio sources which are fainter than about 18 mag in the case of non-stellar objects (galaxies) or about 19.5 mag in the case of stellar objects. These are limits set by optical astronomy. They are set, first, by the fact that the 48 inch Palomar Sky Survey only goes down to 20.0 mag in the red and 21.1 mag in the blue; objects which are fainter than this cannot be identified unless the large reflectors, which typically have very small fields ($\sim 20'$), are used. Second, the difficulties associated with studying very faint objects are very great, because of the increasing difficulty in detecting photons against the natural and man-made noise, even with the use of image-tube techniques; also, only a very limited fraction of observing time on the large telescopes is presently devoted to this work.

Even if one admits that many of the problems can be overcome so that it will eventually be possible to measure the redshifts and do photometry of galaxies down to magnitude 21–23, we still have a problem in finding suitable objects to

observe, because there is no prospect of doing surveys to this level. Sandage and his colleagues³⁰ are now attempting to make identifications by using very accurate radio source positions for which no optical objects can be seen at the level of the 48 inch Palomar Sky Survey. By using special photographic plates and the 200 inch telescope, they have been able to go as faint as magnitude 22 to 23 and find optical objects in some of these positions. Even when such objects are found the problem of measuring redshifts and energy distributions will be very time consuming, and it is not obvious that it will succeed. Those objects which turn out to be QSOs will be useless for this type of investigation. We can use the redshifts of those that turn out to be galaxies, but even then it will be necessary to observe other galaxies which are associated with them to obtain energy distributions which are likely to be sufficiently normal so that they can be used to derive apparent bolometric magnitudes.

Correction for Evolution

A second problem concerns the corrections to the observed energy distribution, based on the view that we are looking at standard galaxies, if they exist. These, when seen at large redshifts, are much younger than the galaxies comparatively close to us, from the viewpoint of stellar evolution. To make such corrections it is necessary to know what types of stars give rise to the bulk of the radiation from elliptical galaxies. A limited amount of work has now been done on the nuclear regions of nearby spirals and on one or two ellipticals in the Virgo cluster. This has led Sandage to conclude that the evolutionary corrections may amount to no more than 0.03 mag per 10^9 years, a value considerably smaller than earlier estimates. Tinsley has argued that theoretical models of galaxies which show a greater degree of evolution, so that the colour correction may be considerably larger, are also compatible with the observations. Most observers are doubtful of her conclusions, but it is far from clear that the present results are conclusive.

At a redshift of $z = 0.2$, the difference in bolometric magnitude of the standard galaxies, m_{bol} , to be expected between $q_0 = 0$ and $q_0 = +1$ is only 0.2 mag, while at $z \approx 0.5$ the difference in m_{bol} between $q_0 = +1$ and $q_0 = -1$ (steady state) is 0.9 mag.

Because the three clusters with redshifts greater than 0.2 appear to lie closest to the $q_0 = +1$ relation, it has been stated for the past 10 years, with different degrees of emphasis, that $q_0 \approx +1$. At the same time it also has sometimes been implied that q_0 is not really known at all. And, in terms of hard data, this latter statement is true. The fact that Baum's measurements have neither been published in detail nor repeated, that we have such a small amount of data anyway, and that the local anisotropy in the velocity field may affect the determination of q_0 all suggest that the numerical value of q_0 is only known to lie between wide limits; so both the exploding models and the steady state models are compatible with the measurements made so far.

Stebbins–Whitford Effect

In 1948 Stebbins and Whitford³² published the first results of their photo-electric measurements of the colours of faint clusters of galaxies. They concluded in this preliminary investigation that after they had put in the usual corrections for the effect of the redshift there appeared to be an excess reddening which increased with increasing redshift. This meant that the distribution of energy radiated by stars in a galaxy is a function of epoch so that an evolutionary effect is present. In the early 1950s, this result was used as observational evidence against the steady state cosmology and there was some discussion in the literature about its significance^{33,34}. But the result was only preliminary and by 1956 Whitford³⁵ no longer believed that it was correct. It was not until 1968 that Oke and Sandage³⁶ put this result to rest. Unfortunately, for many years after the realization by many observers that

the result was spurious, theoreticians were using it as direct evidence against the steady state theory.

Counts of Radio Sources

If any class of extragalactic object is distributed uniformly in space, counts to successively increasing metric distance, and hence to fainter apparent brightnesses, will give numbers which are proportional to the volumes searched. The volumes in Riemannian space depart from the volumes calculated on the Euclidean approximation by factors which are determined by the space curvature. Hubble³⁷ first attempted to count optical galaxies to apply this test but he abandoned the attempt, and more recently Sandage²² has shown that, unless we go to much fainter galaxies than can be observed from the Earth, attempts to distinguish between different cosmological models will not succeed.

The method has, however, been applied extensively to extragalactic radio sources by several groups, who have counted sources at a number of frequencies down to different minimum flux levels. The early work by Ryle and his colleagues^{38,39} gave a log N -log S slope for the brighter sources of -3 , far steeper than the Euclidean value of -1.5 . (If N is the number of sources per steradian brighter than S , and R is the distance to a source, then $N \propto V \propto R^3$, $S \propto R^{-2}$ so that in Euclidean space $N \propto S^{-3/2}$. For distant sources in Riemannian space, $N \propto S^{-3/2} + \mu(S)$ where $\mu(S) \geq 0$ is a function depending on the maximum source distance and the cosmological model.) Quite early on it became clear that the first log N -log S slope obtained in Cambridge was far too steep, and other surveys⁴⁰⁻⁴² all led to the conclusion that at the bright end of the curve the slope is steeper than -1.5 , and has a value somewhere between -1.75 and -1.85 . Following Ryle and his associates it has commonly been argued that this steep slope can only be explained as being due to an excess of sources at great distances. This requires evolution of the source numbers with epoch, and is then evidence against the simple steady state theory. Observations to fainter flux levels show that the slope of the log N -log S plot gradually flattens to a slope of only ~ -0.8 at $S \approx 0.01$ flux units ($1 \text{ f.u.} = 10^{-26} \text{ W m}^{-2} \text{ Hz}^{-1}$). All of these results have been obtained by making surveys at comparatively low frequencies, typical values being 408 MHz and 178 MHz. The work which has been done at higher radio frequencies (2,700 MHz) leads to a log N -log S slope of -1.4 ± 0.1 (ref. 43). If this were the only radio survey no one would have argued that significant departures from a homogenous universe exist. But it has been argued that because of the wide dispersion in spectral indices of the radio sources and their variation with flux density, the slope of the log N -log S curve should decrease with decreasing frequency of observation. The lower frequency survey containing a more homogenous source sample would then be the one most representative of the gross features of the universe. This situation naturally leads to the following unsolved problem: at what observing frequency (radio, microwave, infrared, optical, X-ray) is an unbiased sample most likely to be selected? It would really be preferable to integrate the flux over a suitably large bandwidth and then construct a log N -log S_i curve.

How can we interpret the total slope for the log N -log S curve for sources measured at low frequencies? Ryle and his associates^{44,45} have argued that the data indicate that the universe must be evolving and that there was an excess of sources in an epoch defined by $z \lesssim 2$, but that at even earlier epochs the source density was much lower. The alternative proposal is that either the slope for the bright sources does not depart significantly from the Euclidean value or that its steepness is due to a statistical fluctuation in the number of very bright sources, requiring a deficit of from 5-7 sources per steradian. Hoyle⁴⁶ has argued that the remainder of the curve is directly compatible with the prediction of the steady state model. Ryle and his associates disagree with this, and argue that the whole curve can only be explained by using

Friedmann models with source evolution^{44,45}. The arguments depend on the details of the radio luminosity function and the exact slope of the log N -log S curve for faint sources.

Volume Effect or Luminosity Evolution?

We have optical identifications and redshifts for only a very small fraction of the extragalactic radio sources. But if we restrict ourselves to the 3C (revised) catalogue⁴⁷ containing about 250 of the brightest radio sources ($S \geq 9 \text{ f.u.}$) a total of about 40 per cent of the sources including about forty quasi-stellar objects and sixty-eight galaxies have been identified and redshifts have been obtained. The identification of the QSOs seems to be complete, so the remainder of the sources are thought to be associated with galaxies for which identifications have been made but for which no redshifts have been obtained, or, in a minority of cases, for which there are no optical objects visible on the Palomar Sky Atlas at the radio positions. Plots of radio flux S against redshift have been made both for the QSOs and the galaxies⁴⁸⁻⁵¹ and in each case it has been found that there is no correlation between the apparent flux and the redshift. Thus it seems that for the identified sources the shapes of the log N -log S curves are attributable to a luminosity effect and not to a distance-volume effect. The key to the interpretation of the log N -log S curve for the 3C sources lies in the remaining radio galaxies. If they turn out to have systematically larger redshifts than those already measured, as has been generally assumed because they are optically fainter, then their log N -log S curve is of cosmological significance. If, however, they do not have systematically greater redshifts it must be concluded that the log N -log S curve is dominated by the effects of the radio luminosity function in a comparatively local region. The QSOs do have large redshifts. If these are of cosmological origin then their log N -log S curve and the luminosity volume test⁵² do demonstrate a powerful evolutionary effect. But if the redshifts are not of cosmological origin these analyses give us no evidence of cosmological importance.

A more detailed discussion of the analyses of the counts of sources has been given elsewhere. The limitations set by the source counts on various cosmological models are summarized in Table 1 which is taken from the work of Brecher, Burbidge and Strittmatter⁷⁹. This shows that contrary to popular opinion the counts of radio sources do not rule out a steady state universe.

Quasi-stellar Objects

The QSOs are the only extragalactic objects so far discovered which have very large redshifts.¹ Therefore they are of the greatest importance for cosmology provided that their redshifts are largely due to the expansion of the universe and do not have any other explanation. If the redshifts are of cosmological origin then the luminosity-volume test previously mentioned for the QSOs in the 3CR catalogue, or a similar analysis also carried out by Schmidt⁵³ for a sample of QSOs which are not powerful radio sources, leads to the conclusion that they are not distributed uniformly in space, but that the space density is very strongly dependent on z and hence on epoch. This means that they provide powerful evidence against a simple steady state universe, and if one wishes to preserve the steady state model it must be argued that it is a fluctuating steady state universe in which evolution can occur within an epoch corresponding to redshifts up to ~ 2 . In any case the cut-off of redshifts beyond 2 is of great cosmological significance and some have attempted to relate this to the appropriate epoch in a universe of Lemaitre type^{54,55}.

None of these arguments has any validity unless it can be demonstrated conclusively that the redshifts of the QSOs are of cosmological origin, and this has simply not been achieved. There are a number of arguments that can be adduced for and against the cosmological hypothesis but there is no certain answer. Until this is reached, use of QSOs for cosmological investigation is premature.

Table 1 Limitations Set by Source Counts on Various Cosmological Models

Source	Strict steady state	Fluctuating steady state	Friedmann	Friedmann with source evolution
QSOs				
$z \rightarrow$ Distance	Excluded	Allowed*	Excluded	Allowed†
$z \nrightarrow$ Distance	Allowed	Allowed	Allowed	Allowed
Radio galaxies				
z Known	Allowed	Allowed	Allowed	Allowed
$z \rightarrow$ Distance				
z Unknown	Excluded	Allowed‡	Excluded	Allowed‡
$m_v \rightarrow$ Distance				
m_v Unknown	Excluded§	Allowed	Excluded§	Allowed
$m_v \nrightarrow$ Distance				

* Only if fluctuation to $z \sim 2$ is permitted.† Background limits evolution to $z_{\text{QSO}}^{\text{max}}$ where $2 < z_{\text{QSO}}^{\text{max}} < 4$.‡ Background limits evolution to $z_{\text{RO}}^{\text{max}} \lesssim 0.6$.

§ This case would require a special rather implausible luminosity function to avoid exclusion.

Microwave Background Radiation

As mentioned above, the discovery of microwave background radiation, following the prediction that primaeval fireball radiation generated in a big bang would still be present, was one of the strongest pieces of evidence for such a beginning. Is the information that we now have on the microwave background still compatible with this interpretation? Observations made at wavelengths between a few millimetres and about 20 cm fit quite well on a black-body curve with a temperature of about 2.7 K. The energy density of the radiation is then about 4×10^{-13} erg cm⁻³. If the black-body form is confirmed, this will be the strongest evidence in favour of an evolving universe. But it is not yet certain whether or not the spectrum is really of black-body form, because the wavelength range directly observed, and not in dispute, is the Rayleigh-Jeans part of the curve where $I(\nu) \propto \nu^2$.

The observations which have been made either directly or indirectly close to and just beyond the peak of the black-body curve are in conflict unless it is argued that some very strong line radiations are present. The indirect observation of the radiation field in the galactic plane which is exciting the CN, CH and CH⁺ interstellar molecules suggests that the temperature is less than ~8 K at 9.36 mm, less than 5 K at 0.56 mm, less than ~4 K at 1.3 mm, and it appears to be about 2.8 K at 2.64 mm. Direct observation from rockets and a balloon above the Earth's atmosphere at wavelengths in the range 0.4–1.3 mm gives a flux which corresponds to a black-body temperature of about 6 K (refs. 56–58). It is possible that the radiation detected by these observers is concentrated in a line which could be produced in a geocorona, in the Galaxy, or have a cosmological origin. The interpretation of the results is unclear at present.

If it turns out that the microwave radiation field departs strongly from the black-body form, then either extra components are present, or the radiation has a quite different origin. In both cases the answer must be that the excess radiation, or the total flux, ultimately arises in discrete sources. A number of investigations have been made in which these possibilities have been explored. The sources are most likely to be the nuclei of galaxies which radiate largely at infrared and microwave frequencies.

It was the discovery of the microwave background radiation that led to a large number of theoretical investigations concerned with the early history of the universe. What should be made clear, however, is that the existence of this radiation is very powerful evidence for an initially dense state, provided that the radiation has a black-body spectrum as is predicted in the simple theory. If it turns out that it departs strongly from the black-body shape, many attempts will undoubtedly be made to explain it in terms of a primordial radiation field plus large contributions from discrete sources, or even a more

complex big bang. But this must not obscure the important point that the absence of a black-body form would mean that the prediction has not been fulfilled and the strong and direct evidence for an initial dense state will have disappeared.

This concludes my discussion of direct observational evidence bearing on whether or not the universe is evolving and began in a dense state. I believe that if one attempts to evaluate this evidence objectively there is still no really conclusive evidence in favour of such a universe.

Nucleosynthesis and Origin of Galaxies

Another approach can be taken in looking for evidence which bears on the question of whether or not the universe began in a dense state. The bulk of the mass that we see is condensed in the form of stars and galaxies with a wide range of masses and angular momenta. This matter is not distributed uniformly in space but is condensed into dense aggregates. Moreover, it has a complicated chemical composition. The major problems of cosmogony are therefore to understand the origin of the elements, and the formation of discrete objects—galaxies and other compact massive objects.

Does the existence of these features of the universe suggest that their origin is closely connected with a big bang?

I shall consider first the problem of nucleosynthesis. Early work on this problem made little progress because knowledge of nuclear physics was still comparatively primitive. It was Gamow, Alpher and Herman^{59–63} who first made a serious attempt to explain the synthesis of the heavier elements from an initial dense cloud of baryons, leptons and radiation in a hot big bang. It is well known that this attempt failed due to the difficulty associated with building beyond mass 4. In the 1950s a serious attempt was made to understand the relative abundances of all of the isotopes on the assumption that they have nearly all been synthesized in the interior of stars⁶⁴. This theory has been, in large part, highly successful and is generally accepted. It was originally thought that the fact that heavy elements are synthesized from lighter ones in stars meant that there is a slow build-up of heavy elements as a galaxy grows older, and that the relative abundance would be correlated with the ages of the stars. It is now clear that while this effect is operating, a large amount of nucleosynthesis takes place early in the life of a galaxy and the processes going on steadily throughout its life are in some ways of lesser importance.

From the point of view of cosmology the most important result is that very detailed calculations made by Wagoner, Fowler and Hoyle⁶⁵ have demonstrated conclusively that the bulk of the elements cannot have been made in a big bang. The only important element, from the point of view of nuclear physics, that could have been synthesized in a big bang is helium. Thus soon after the discovery of the microwave background radiation there was a serious attempt to argue that the helium that is detected in stars and gaseous nebulae in our own galaxy and in other galaxies is primordial helium made in a big bang. The argument appeared plausible because, on the one hand, it was suggested that the helium/hydrogen ratio is the same in different celestial objects (stars, gaseous nebulae in our own galaxy, and a few gaseous nebulae in one or two nearby galaxies) and is 25–30 per cent by mass, while, on the other hand, the fractional amount of helium relative to hydrogen that is made in a big bang is close to about 27 per cent. This simple argument broke down, however, as soon as detailed studies of different aspects of the problem were made. First of all, it was pointed out that there appear to be stars which contain very little helium so that perhaps there is no universal helium abundance. While this argument is still open, detailed studies of big bang models led to the conclusion that there are ways of making very little helium in such an initial process^{66–68}. Fowler has recently emphasized that the helium-hydrogen ratio is determined completely by the choice of baryon and lepton numbers in a big bang⁶⁹.

which has nothing to do with the theory of general relativity, but is tied only to initial conditions which lie outside the known laws of physics. Thus, while the helium abundance in the universe might be related to the occurrence of a big bang, there is no direct evidence that the two are related, and there are a number of alternative ways of understanding the large helium abundance found in some stars and galaxies. For example, it appears that helium could have been synthesized in massive objects evolving in the nuclear regions of galaxies.

Finally, we come to the vexed question of the origin of galaxies. As the big bang bandwagon has gained momentum, an increasing number of investigations have been carried out in which attempts are made to explain the condensation of dense objects from an initial cloud of matter and radiation which is expanding. It has been known for many years that this is very difficult to understand⁷⁰⁻⁷², and the investigations have now reached the point where it is generally accepted that the existence of dense objects cannot be understood unless very large density fluctuations in a highly turbulent medium, or otherwise, are invoked in the first place⁷³⁻⁷⁵. There is again no physical understanding of the situation; it is a condition which is put in, in a hypothetical state, to explain a major property of the universe. Thus these "theories" amount to nothing more than the statement that protogalaxies have a cosmological origin, and their origin cannot be understood any better than can the original baryons and leptons in an evolving universe. At this level therefore these theories are in no better shape than the apparently much more radical views of Ambartsumian^{76,77} or Hoyle and Narlikar⁷⁸ who believe that the origin of galaxies is tied closely to the properties of the dense active nuclei and that perhaps the initial states of galaxies were high density states. Hoyle and Narlikar have even dared to suggest that perhaps "new" physics is involved. What the debate largely reduces to is that on one side there is an apparent belief by the majority that creation in the distant past is acceptable, but that creation at recent epochs is unthinkable, while on the other it is thought that there is little to choose between the two points of view, and that only with much more original work can we hope to resolve the issue. It is clear that the evolving universe concept gains no support from attempts to understand either the origin of the elements or the origin of galaxies.

Was there really a big bang? I believe that the answer clearly must be that we do not know, and that if we are ever to find an answer much more effort must be devoted to cosmological tests, with a much more openminded approach, and that much more original thinking must be done to attempt to explain the large amount of observational material, and not only that material that can be used in a narrow sense to fit preconceived ideas. Probably the best argument in favour of a beginning is the general result that the ages of many stars in our galaxy are approximately equal to H^{-1} . Probably the strongest argument against a big bang is that when we come to the universe in total and the large number of complex condensed objects in it, the theory is able to explain so little.

- ¹ North, J. D., *The Measure of the Universe* (Clarendon Press, Oxford, 1965).
- ² Dingle, H., *Observatory*, **73**, 42 (1953).
- ³ Novikov, I. D., and Zeldovich, Y. B., *Ann. Rev. Astron. Astrophys.*, **5**, 627 (1967).
- ⁴ Lemaitre, G., *Ann. Soc. Sci. Bruxelles*, **47**, A49 (1927).
- ⁵ Lemaitre, G., *Mon. Not. Roy. Astron. Soc.*, **91**, 483 (1931).
- ⁶ Friedmann, A., *Z. Physik*, **10**, 377 (1922).
- ⁷ Humason, M. L., *Proc. US Nat. Acad. Sci.*, **15**, 167 (1929).
- ⁸ Hubble, E. P., *Proc. US Nat. Acad. Sci.*, **15**, 168 (1929).
- ⁹ Bondi, H., and Gold, T., *Mon. Not. Roy. Astron. Soc.*, **108**, 252 (1948).
- ¹⁰ Hoyle, F., *Mon. Not. Roy. Astron. Soc.*, **108**, 372 (1948).
- ¹¹ Peach, J. V., *Intern. Astron. Union Symp.*, No. 44, Uppsala (1970).
- ¹² Sandage, A. R., *Mayall Symp.*, Tucson, May (1971).
- ¹³ Iben, I., and Faulkner, J., *Astrophys. J.*, **153**, 101 (1968).
- ¹⁴ Iben, I., and Rood, R., *Astrophys. J.*, **161**, 587 (1970).
- ¹⁵ Demarque, P., Mengel, J. G., and Aizenman, M., *Astrophys. J.*, **163**, 37 (1971).
- ¹⁶ Fowler, W. A., *Rutherford Symposium Volume* (1968).

- ¹⁷ Alpher, R. A., and Herman, R., *Phys. Rev.*, **75**, 1089 (1949).
- ¹⁸ Gamow, G., *Kgl. Danske Videnskab Selskab Mat. Fys. Medd.*, **27**, No. 10 (1953).
- ¹⁹ Penzias, A., and Wilson, R., *Astrophys. J.*, **142**, 419 (1965).
- ²⁰ Humason, M. L., Mayall, N. U., and Sandage, A. R., *Astron. J.*, **61**, 97 (1956).
- ²¹ Sandage, A. R., *Astrophys. J.*, **127**, 513 (1958).
- ²² Sandage, A. R., *Astrophys. J.*, **133**, 355 (1961).
- ²³ Sandage, A. R., *Proc. Course 35, International School of Physics Enrico Fermi*, 10 (Academic Press, New York, 1966).
- ²⁴ Sandage, A. R., *Observatory*, **92**, 91 (1968).
- ²⁵ Sandage, A. R., *Phys. Today*, **23**, 34 (1970).
- ²⁶ Peach, J. V., *Astrophys. J.*, **159**, 753 (1970).
- ²⁷ Baum, W. A., *Astron. J.*, **62**, 6 (1957); Baum, W. A., *Intern. Astron. Union Symp.*, No. 15 (edit. by McVittie, G. C.), 390 (Macmillan, New York, 1961).
- ²⁸ Minkowski, R., *Astrophys. J.*, **132**, 908 (1960).
- ²⁹ Burbidge, G. R., *Ann. Rev. Astron. Astrophys.*, **8**, 369 (1970).
- ³⁰ Sandage, A. R., *Pont. Acad. Scient. Scripta. Varia*, **35**, *Nuclei of Galaxies*, Vatican, 601, (1971).
- ³¹ Kristian, J., and Sandage, A. R., *Astrophys. J.*, **162**, 391 (1970).
- ³² Stebbins, J., and Whitford, A. E., *Astrophys. J.*, **108**, 413 (1948).
- ³³ Bondi, H., Gold, T., and Sciama, D., *Astrophys. J.*, **120**, 597 (1954).
- ³⁴ Whitford, A. E., *Astrophys. J.*, **120**, 599 (1954).
- ³⁵ Whitford, A. E., *Astron. J.*, **61**, 353 (1956).
- ³⁶ Oke, J. B., and Sandage, A. R., *Astrophys. J.*, **154**, 21 (1968).
- ³⁷ Hubble, E., *The Realm of the Nebulae* (Oxford University Press, 1936).
- ³⁸ Ryle, M., *Observatory*, **75**, 137 (1955).
- ³⁹ Ryle, M., and Scheuer, P. A. G., *Proc. Roy. Soc.*, **A**, **230**, 448 (1955).
- ⁴⁰ Mills, B. Y., *Paris Symp. on Radio Astronomy* (edit. by Bracewell, R. N.) (Stanford, 1959).
- ⁴¹ Day, G. A., Shimmins, A. T., Ekers, R. J., and Cole, D. J., *Austral. J. Phys.*, **9**, 35 (1966).
- ⁴² Ryle, M., and Scott, P. F., *Mon. Not. Roy. Astron. Soc.*, **122**, 389 (1961).
- ⁴³ Shimmins, A. T., Bolton, J., and Wall, J. V., *Nature*, **217**, 818 (1968).
- ⁴⁴ Ryle, M., *Ann. Rev. Astron. Astrophys.*, **6**, 249 (1968).
- ⁴⁵ Longair, M., *Mon. Not. Roy. Astron. Soc.*, **133**, 421 (1966).
- ⁴⁶ Hoyle, F., *Proc. Roy. Soc.*, **A**, **308**, 1 (1969).
- ⁴⁷ Bennett, A. S., *Mem. Roy. Astron. Soc.*, **68**, 163 (1962).
- ⁴⁸ Hoyle, F., and Burbidge, G. R., *Nature*, **210**, 1346 (1966).
- ⁴⁹ Longair, M., and Scheuer, P. A. G., *Nature*, **215**, 919 (1967).
- ⁵⁰ Hoyle, F., and Burbidge, G. R., *Nature*, **227**, 359 (1970).
- ⁵¹ Bolton, J. G., *Astron. J.*, **74**, 131 (1969).
- ⁵² Schmidt, M., *Astrophys. J.*, **151**, 393 (1968).
- ⁵³ Schmidt, M., *Astron. J.*, **162**, 371 (1970).
- ⁵⁴ Petrosian, V., Salpeter, E. E., and Szekeres, P., *Astrophys. J.*, **147**, 1222 (1967).
- ⁵⁵ Shklovsky, I. S., *Astrophys. J. Lett.*, **150**, L1 (1967).
- ⁵⁶ Shivanandan, K., Houck, J. R., and Harwit, M., *Phys. Rev. Lett.*, **21**, 1460 (1968).
- ⁵⁷ Houck, J. R., and Harwit, M., *Astrophys. J.*, **157**, L45 (1969).
- ⁵⁸ Muehlner, D., and Weiss, R., *Phys. Rev. Lett.*, **24**, 742 (1970).
- ⁵⁹ Gamow, G., *Phys. Rev.*, **70**, 572 (1946).
- ⁶⁰ Gamow, G., *Phys. Rev.*, **74**, 505 (1948).
- ⁶¹ Alpher, R. A., *Phys. Rev.*, **74**, 1577 (1948).
- ⁶² Alpher, R. A., and Herman, R. C., *Phys. Rev.*, **74**, 1737 (1948).
- ⁶³ Alpher, R. A., and Herman, R. C., *Rev. Mod. Phys.*, **22**, 153 (1950).
- ⁶⁴ Burbidge, E. M., Burbidge, G. R., Fowler, W. A., and Hoyle, F., *Rev. Mod. Phys.*, **29**, 547 (1957).
- ⁶⁵ Wagoner, R., Fowler, W. A., and Hoyle, F., *Astrophys. J.*, **148**, 3 (1967).
- ⁶⁶ Hawking, S. W., and Taylor, R. J., *Nature*, **209**, 1278 (1966).
- ⁶⁷ Thorne, K., *Astrophys. J.*, **148**, 51 (1967).
- ⁶⁸ Dicke, R., *Astrophys. J.*, **152**, 1 (1968).
- ⁶⁹ Fowler, W. A., *Comments on Astrophys. and Space Phys.*, **2**, 134 (1970).
- ⁷⁰ Lifshitz, E., *J. Phys. (USSR)*, **10**, 116 (1946).
- ⁷¹ Bonnor, W. B., *Z. Astrophys.*, **39**, 143 (1956).
- ⁷² Lifshitz, E., and Khalatnikov, I. M., *Adv. Phys.*, **12**, 185 (1963).
- ⁷³ Harrison, E. R., *Mon. Not. Roy. Astron. Soc.*, **148**, 119 (1970).
- ⁷⁴ Ozernoy, L. M., and Chernin, A. D., *Astron. Zh.*, **44**, 1131 (Sov. Astron., **11**, 907) (1967).
- ⁷⁵ Ozernoy, L. M., and Chernin, A. D., *Astron. Zh.*, **45**, 1137 (Sov. Astron., **12**, 901) (1968).
- ⁷⁶ Ambartsumian, V. A., *Solvay Conf. on Structure and Evolution of the Universe*, Brussels (edit. by Stoops, R.), 241 (Univ. of Brussels, 1958).
- ⁷⁷ Ambartsumian, V. A., *Structure and Evolution of Galaxies*, *Proc. Thirteenth Conf. Phys.*, Univ. Brussels, 1 (Wiley, Interscience, 1965).
- ⁷⁸ Hoyle, F., and Narlikar, J. V., *Proc. Roy. Soc.*, **A**, **290**, 177 (1966).
- ⁷⁹ Brecher, K., Burbidge, G., and Strittmatter, P., *Comments Astrophys and Space Physics* (in the press).

On the Nature of Mass

F. HOYLE & J. V. NARLIKAR

Institute of Theoretical Astronomy, University of Cambridge

The increasing number of observations of discrepant redshifts means that no longer can these be passed off as chance juxtapositions. A possible explanation of the data is given here in terms of a theory that incorporates a gravitational "constant" that is decreasing with time.

EVIDENCE concerning "discrepant redshifts" has accumulated fairly rapidly in recent months. The case where it is hardest to deny the evidence was reported at the beginning of the present year by Arp¹. The galaxy NGC 7603 was found to be connected to an appendage by two arms which intersect at the appendage $z=0.056$. In the customary Doppler interpretation of the redshift this difference exceeds $8,000 \text{ km s}^{-1}$. A redshift difference of the same order was found many years ago² by Minkowski, Humason and Zwicky in "Zwicky's triple system". The galaxies in this system are also connected by a bridge.

If it is accepted that a discrepancy of $\sim 8,000 \text{ km s}^{-1}$ exists for NGC 7603, then there seems no good reason for denying the two cases with discrepancies $\sim 20,000 \text{ km s}^{-1}$ obtained by Sargent (ref. 3, and W. L. W. Sargent's contribution to the Vatican Conference, 1970). Both the latter cases are compact groups of galaxies in which one member has the discrepant redshift. Less than ten such groups have been examined. The probability of a more distant galaxy being projected against a nearer cluster by chance is small. Arp (private communication) has obtained what appears to be a jet connecting the quasi-stellar object Makarian 205, $z \simeq 0.07$, to the galaxy NGC 4319. The discrepancy is again of order $20,000 \text{ km s}^{-1}$.

Three cases have been observed of a quasi-stellar object apparently associated with a small compact cluster of faint galaxies. In one case Gunn has reported⁴ the same redshift for the QSO as for one of the galaxies of the cluster. In the other two cases there appear to be large redshift differences (Hazard, Jauncey, Sargent and Gunn, private communication).

In a recent preprint Burbidge, Burbidge, Solomon and Strittmatter have drawn attention to a bridge which appears on the Palomar sky prints to connect the quasi-stellar object PHL 1226 to the bright galaxy IC 1746. The redshift of PGL 1226 is $z=0.404$, whereas that of the galaxy has not yet been measured, but must be small because of its brightness.

Because all objects are projected on the sky there must be some apparently peculiar juxtapositions of objects that are really at very different distances. This has hitherto led to a situation in which all but a very few astronomers have dismissed apparent redshift discrepancies as simply unusual projection effects. Yet there has to be a point of balance in

one's judgment as to the extent of the chance juxtapositions one is willing to accept. One of us was aware at an early stage of the case of Zwicky's system (Zwicky, private communication) but it seemed that, in spite of the bridge, there might be a very peculiar chance juxtaposition. But the case of VV 172³ formed the point of balance and NGC 7603 has turned the balance.

If a highly convincing theory of discrepant redshifts were available then we think there is little doubt that the data would today be considered reasonably clear-cut. It has been the absence of such a theory that has caused most astronomers to prefer to believe in unusual projection effects. If, as seems very possible, the accumulation of data forces us over a watershed (not only in our thinking but in the history of astronomy) it will clearly become necessary to arrive at a theory of discrepant redshifts. We wish to emphasize the need for a thoroughly radical assessment of the problem, considering it unlikely that a satisfactory theory will be achieved by a small change in our concepts. Explicitly, we do not think discrepant redshifts will be explained adequately either as simple Doppler peculiar motions or as excess reddening due to gravitation.

In the following discussion we shall present the usual Friedmann cosmological models and the Einstein de Sitter case in particular, from an unusual point of view. Mathematically, this point of view is at first entirely equivalent to the usual formulation. Then at a later stage we shall arrive at a possible major shift from the usual theory. The balance between the usual theory and an entirely different view both of physics and astrophysics hinges on a question raised by Dirac⁵ in 1937.

It is well known that when c and $\hbar$ are set equal to unity, only a single dimensionality is needed for the whole of physics. We take this to be length and we denote the unit by L . Every quantity has a dimensional form L^n , for example, mass $\sim L^{-1}$, frequency $\sim L^{-1}$, charge $\sim L^0$, magnetic field $\sim L^{-2}$ and the gravitational 'constant' has dimensionality L^2 .

Every observation is concerned with a dimensionless number, so that every observation is concerned with a product of quantities such that the dimensional dependencies on L cancel to zero. So far as experimental physics and engineering are concerned it is possible to convert a quantity of dimensionality L^n into a quantity of dimensionality L^m by means of a linear device provided $n=m$. Non-linear devices are needed if $n \neq m$. This property makes it comparatively easy to see what kind of physical device is needed to relate one quantity to another. There is no doubt that physics and engineering are made quite unnecessarily complicated by the current practice of using multiunit systems.

Because observed quantities are dimensionless they are unaffected by a scale change in the length unit. But what if we elect to change the length unit in accordance with a well behaved function $\Omega(x)$ (no zeros of Ω as well as no infinities) which varies with the position of the space-time point X ? Nothing should be changed provided that first, we always combine physical quantities at the same X , and, second,

every physical quantity is scaled according to its dimensionality—a quantity of dimensionality L^n being scaled by Ω^n .

Physical theories with these properties are said to be conformally invariant. This is a different kind of invariance from the coordinate invariance of relativity. Not all physical theories, however, are conformally invariant. Maxwell's theory is conformally invariant. The derivative terms of Dirac's equation are conformally invariant, but the mass term is not. Einstein's theory is not conformally invariant and physical theories cease to be conformally invariant whenever "mass" is involved. This is because the second property is then not satisfied. Clearly, we cannot expect classical dynamics to be conformally invariant unless the classical action

$$S = - \int m \, da \quad (1)$$

of particle a is invariant. Since the element da of the path of the particle becomes multiplied by Ω we would require the mass m to be multiplied by Ω^{-1} , which indeed would accord with the dimensionality of mass, $\sim L^{-1}$. But Ω can now be different at different space-time points, so that in general the mass of the particle will not be the same for all points on its path. Plainly then, we cannot expect to arrive at a conformally invariant system of dynamics so long as the mass of a particle is considered a fixed quantity belonging autonomously to the particle.

We have been concerned^{6,7} since 1964 in developing a conformally invariant theory of dynamics and of gravitation. The central idea of this work is that the mass of a particle m say, is given by

$$m(X) = \lambda M(X) \quad (2)$$

where λ is a coupling constant belonging to the particle itself and $M(X)$ is a "mass field" generated by all the particles in the universe. In a very general way we can write

$$M(X) = \sum_b \int P(X, B) \, db \quad (3)$$

the summation being over all particles b , the line integral being over the world line of b , and $P(X, B)$ being a propagation function from point B at db to the field point X . To avoid self-action problems in the classical theory it is convenient to add the refinement that in determining the mass of any particular particle a this particle is omitted in the summation with respect to b . A comparable situation occurs in classical electrodynamics.

Mathematically we require that when all lengths on the right hand side of (3) are changed by the function Ω the resulting mass field is multiplied by $\Omega^{-1}(X)$. This requirement is met by putting

$$P(X, B) = \lambda \tilde{G}(X, B) \quad (4)$$

where $\tilde{G}(X, B)$ is the time-symmetric elementary solution of

$$\square_x \tilde{G}(X, B) + 1/6 R(X) \tilde{G}(X, B) = [-g(B)]^{-1} \delta_4(X, B) \quad (5)$$

in which $\square_x$ is the d'Alembertian operator, R the scalar Riemannian curvature and δ_4 the 4-dimensional Dirac delta-function. So far as we are aware the wave-equation (5) was first considered with respect to its conformal properties by Penrose.

The action of particle a is now

$$-\lambda^2 \int da \cdot \sum_{b \neq a} \int \tilde{G}(A, B) \, db \quad (6)$$

and the total action of all particles is given by

$$S = -\lambda^2 \sum_a \sum_b \iint \tilde{G}(A, B) \, da \, db \quad (7)$$

Subtleties arise for a universe with infinitely many particles concerning the finiteness of (7), but the situation in this respect is not really different from the usual theory.

Enough has already been done to make the Dirac equation conformally invariant, including the mass term. To obtain conformally invariant gravitational equations we work with the usual condition that S be stationary for small changes of the metric tensor, $\delta S = 0$ for $g_{ik} \rightarrow g_{ik} + \delta g_{ik}$. The resulting equations are different and more complicated than those of Einstein. For cosmology however, and for all problems concerned with weak gravitational fields, the equations can be reduced to the Einstein form with an exceedingly high degree of approximation. The situation in regard to the classical tests of general relativity—the bending of light, the perihelion of Mercury—is therefore exactly the same as with Einstein. To effect the reduction to Einstein's equations one chooses the conformal frame so that $M(X)$ becomes independent of X . If, to begin with, $M(X)$ varies with X we simply choose $\Omega \propto M$. Particle masses are independent of X in this particular conformal frame, just as we normally take them to be.

The gravitational 'constant' comes out to be given by

$$G = \frac{3}{4\pi M^2} \quad (8)$$

Again in the special conformal frame in which M is constant, G is indeed a constant. It is interesting that G is necessarily positive, so that weak gravitational fields are necessarily attractive.

It should be added that the gravitational equations do not reduce to the Einstein form in certain cases where the field is strong. The discussion of "black holes" is greatly changed, for example.

Because the gravitational theory is conformally invariant we can now effect a major simplification in discussing the isotropic homogeneous models usually employed in cosmology. These models are usually discussed with respect to the Robertson-Walker form of the line element

$$ds^2 = d\tau^2 - Q^2(\tau) \left[\frac{d\rho^2}{1 - k\rho^2} + \rho^2 (d\theta^2 + \sin^2\theta \, d\phi^2) \right] \quad (9)$$

Here τ is cosmic time and $Q(\tau)$ is the expansion function. The coordinates ρ, θ, ϕ are spherical polars for an observer at the origin. The constant k can be zero or ± 1 . $k=0$ gives the Einstein de Sitter model, $k=+1$ is closed, $k=-1$ hyperbolic. It is well known that a conformal transformation function Ω can be found that changes the line element (9) into flat Minkowski space.

It is not usually possible to take advantage of this geometrical simplification because the physical theory is not conformally invariant. Here, however, the physics is conformally invariant and we can transform to Minkowski space. When we have done so the theory is as good as it was in the Robertson-Walker representation. It turns out that while the $k=0$ case is still homogeneous in the Minkowski space

$$dt^2 - dr^2 - r^2 (d\theta^2 + \sin^2\theta \, d\phi^2) \quad (10)$$

the cases $k = \pm 1$ are simply localized clouds of no special interest.

The full Minkowski space, $t < 0$ as well as $t > 0$, can be used. There is, however, nothing corresponding to both halves of Minkowski space in the Robertson-Walker representation. The reason for this emerges when we make the inverse transformation from (10) to (9). The inverse transformation has a singularity at $t=0$, and this introduces a singularity at $\tau=0$ in (9). This singularity, which now emerges as a mathematical construct, is usually interpreted as the 'origin' of the universe. There need be no 'origin' in the Minkowski representation.

Taking the particle masses to be constant in the Robertson-Walker frame, and applying Ω^{-1} to the mass, Ω being chosen for transformation from (9) with $k=0$ to (10), we obtain

$$m = (\text{constant}) \cdot t^2 \quad (11)$$

The particle masses are a quadratic function of cosmic time t in the Minkowski representation. Even though space is now flat this leads to a redshift effect. Remembering that atomic sources emit frequencies proportional to m , we see that such a source at distance r from an observer at $r=0$ would be judged to have a redshift given by

$$1+z = \left(\frac{t}{t-r} \right)^2 \quad (12)$$

the time of observation being t and the time of emission $t-r$.

A standard galaxy may be considered to have a specified number of stars each containing a specified number of particles arranged in the same way with respect to chemical composition. The luminosity of a star has dimensionality L^{-2} . This is proportional to m^2 , and therefore by (11) proportional to the fourth power of the cosmic time. The bolometric flux from such a standard galaxy observed at time t is therefore

$$\frac{L(t-r)}{4\pi r^2} \propto \frac{(t-r)^4}{r^2} \quad (13)$$

where $L(t-r)$ is the luminosity of the galaxy at the time $t-r$ of emission of the observed radiation. Eliminating r from (13) with the aid of (12) we find that the bolometric flux depends on z according to

$$\frac{1}{1+z} \cdot \frac{1}{[\sqrt{1+z}-1]^2} \quad (14)$$

which is the classic Hubble redshift-magnitude relation for the Einstein de Sitter model. The reader will note how much more easily we have arrived at (14) than in the usual treatment.

The present treatment is still deficient in one important respect. The form of equation (11) for the particle mass was obtained for the Minkowski representation by assuming constant mass in the Robertson-Walker representation. To make the argument complete it is essential to show that (11) follows from an explicit evaluation of (2), (3), (4), (5). This evaluation turns out to involve interesting problems which will not be considered here, since details have been given elsewhere⁸. It is convenient to choose the length unit L so that the particle density is everywhere unity. We then obtain from (2), (3), (4), (5),

$$m = \frac{1}{2} \lambda^2 t^2 L^{-3} \quad (15)$$

which has the required quadratic dependence on t .

Expanding (12) for $r/t \ll 1$ we obtain

$$z \simeq 2r/t \quad (16)$$

Since the Hubble constant H is defined by $z = Hr$ for small r , we have

$$t = 2H^{-1} \quad (17)$$

for the present epoch. Empirically $H^{-1} \simeq 2 \times 10^{28}$ cm. Empirically also, if we consider the particles to be hydrogen atoms, the average particle density n is given by

$$n = L^{-3} \simeq 2.5 \times 10^{-6} \text{ cm}^{-3} \quad (18)$$

From (17), (18), and $H^{-1} = 2 \times 10^{28}$ cm,

$$nt^3 \simeq 10^{80} \quad (19)$$

This cosmological number is dimensionless, and is the number of particles contributing to the mass field $M(X)$ at the present epoch.

The dimensionless number λ^2 has no effect on gravitation, since it multiplies the action (7), and therefore disappears entirely from $\delta S = 0$, $g_{ik} \rightarrow g_{ik} + \delta g_{ik}$. But the value of λ^2 has a profound effect on the behaviour of individual particles when electromagnetic terms are added to the action. The value of

λ^2 can be obtained empirically in the following way. Applying (15) to protons,

$$m_p = \frac{1}{2} \lambda^2 t^2 L^{-3} = 2 \lambda^2 H^{-2} L^{-3} \quad (20)$$

for the present epoch. The empirical values (18) for L^{-3} and $H^{-1} \simeq 2 \times 10^{28}$ cm then lead to

$$m_p \simeq 2 \times 10^{51} \lambda^2 \text{ cm}^{-1} \quad (21)$$

The laboratory value of the Compton wavelength of the proton m_p^{-1} is $\sim 10^{-14}$ cm, so that empirically $m_p \simeq 10^{14} \text{ cm}^{-1}$, and

$$\lambda^2 \simeq 5 \times 10^{-38} \quad (22)$$

is therefore required by atomicity.

The remarkable fact now emerges that

$$\lambda^2 (nt^3)^{\frac{1}{2}} = O(1) \quad (23)$$

On the left λ^2 apparently belongs autonomously to the particle, whereas $(nt^3)^{\frac{1}{2}}$ is a cosmological number. Apparently also λ^2 and n are independent of t , so that (23) would seem only a coincidence of the present epoch. This is the usual view.

At this point we reach our watershed. How if we require that (23) shall apply at all t ? Then λ^2 or n , or both, must depend on t . This was the proposal of Dirac⁵. It is clear that (23) can be maintained by $n\lambda^4 \propto t^{-3}$, which allows many possible combinations of n , λ^2 . Here we discuss a possibility which involves no change in the mass field $M(X)$. This possibility is similar in some respects to the case discussed by Dirac but is different in other respects. According to (2), (3), (4),

$$M(X) = \lambda \sum_b \int \tilde{G}(X, B) db \quad (24)$$

from which it is not hard to see that provided the product λn is independent of t there will be no change in $M(X)$. From

$$\lambda \propto n^{-1}, n\lambda^4 \propto t^{-3} \quad (25)$$

we then get

$$\lambda \propto t^{-1}, n \propto t \quad (26)$$

Moreover, the total action (7) is the same as before. It is as if more particles are being created by arranging that the masses of individual particles, $m = \lambda M$, shall increase linearly with t , instead of as t^2 . In so far as we discuss the gravitational aspects of cosmology in terms of a smooth fluid, instead of in terms of discrete particles, nothing is changed. The interaction of each particle with the mass field M weakens as t^{-1} , but there are more particles because of $n \propto t$, and the total interaction of all particles remains the same.

An important by-product of these considerations is that the universe now becomes a perfect absorber in the sense of Wheeler and Feynman. The original Einstein de Sitter model did not meet the perfect absorber condition, which in our view is a fatal flaw of this model. The present model is not the steady state model, but by moving some way towards the steady state model, by permitting $n \propto t$, we have managed to meet an essential electrodynamic requirement.

It is possible to make a conformal transformation $\Omega \propto t$ from the Minkowski representation to a representation in which individual particle masses are constant, as we consider them to be in all terrestrial and astrophysical problems. The line

$$ds^2 = dT^2 - 2T[dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)] \quad (27)$$

which is the same as the line element of a radiation-dominated Friedmann model. The present model, however, is not radiation-dominated and it has a gravitational 'constant' that varies with time. In this representation the mass field M is proportional to $T^{\frac{1}{2}}$ and G , given by (8), is proportional to T^{-1} . Such a variation would have a profound effect on astrophysics and geophysics.

Stars would be much brighter at early T than we usually suppose them to be. There are advantages to be gained from such behaviour, but $G \propto T^{-1}$ may be too drastic⁹, for example, in its effect on the past luminosity of the Sun. On the other hand, there are aspects of geophysics that seem as if they would be greatly helped by this kind of dependence. Steadily weakening gravity would gradually release the interior of the Earth from compression. It can be calculated⁸ that the radius of the Earth would be required to increase at about 10 km per 10^8 yr. There is no possibility of this expansion being resisted by the crust, which must be cracked open repeatedly to make way for new surface material. At all times there would be an excess upward force on the crust at the limit of its strength. The possibility of large horizontal pressure differences, of order 10^9 dyne cm^{-2} , also exists, provided in particular regions that excess pressure is conveyed to the immediate subsurface by fluid material.

We are reminded in this connexion of the old controversy concerning continental drift. Our impression is that, while modern evidence shows unequivocally that drift actually takes place, the early calculations demonstrating the need for exceedingly large forces, really remain valid. If this is so, we would be inclined to think that some such behaviour of G as is given by our model becomes essential for an understanding of the geophysical evidence.

It may be added that Shapiro, Smith, Ash, Ingalls and Petten-gill¹⁰ have recently placed an observational upper limit of 4×10^{-10} yr^{-1} on $\dot{G}/G$. The variation expected here, with a Hubble constant of 5×10^{-11} yr^{-1} , would be 10^{-10} yr^{-1} .

We turn finally to the problem of how we should interpret $\lambda \propto t^{-1}$ in relation to the redshift problems discussed at the

outset. The proportionalities $n \propto t$, $\lambda \propto t^{-1}$ maintain cosmological homogeneity. We must contemplate, however, that these proportionalities represent the smoothed-out effect of changes that could possess spatial irregularities. There could be spatially adjacent variations of λ . This we feel to be the kind of concept needed to come to grips with the problem of redshift anomalies. The smoothed-out homogeneous behaviour of the particle masses gives the normal cosmological redshift, obeying the usual Hubble relation. Local variations give anomalies. A more rapid decrease locally of λ is required in order that the anomalies be in the sense of an increased redshift. One can speculate that such variations occur as a consequence of the physical conditions in regions of strong gravitational fields. These are characterized by the condition that local particles make a contribution to the total mass field M that is comparable with the contribution of particles at cosmological distances. Such localities may be able to produce their own environmental conditions, leading to the local variations of λ .

Received July 27, 1971.

¹ Arp, H. C., *Astrophys. Lett.*, **7**, 221 (1971).

² Zwicky, F., *Naturwissenschaften*, **29**, 344 (1956).

³ Sargent, W. L. W., *Astrophys. J. Lett.*, **153**, L135 (1968).

⁴ Gunn, J., *Astrophys. J. Lett.*, **164**, L113 (1971).

⁵ Dirac, P. A. M., *Nature*, **139**, 323 (1937).

⁶ Hoyle, F., and Narlikar, J. V., *Proc. Roy. Soc., A*, **294**, 138 (1966).

⁷ Hoyle, F., and Narlikar, J. V., *Ann. Phys.*, **62**, 44 (1971).

⁸ Hoyle, F., and Narlikar, J. V., *Mon. Not. Roy. Astron. Soc.* (in the press).

⁹ Pochoda, P., and Schwarzschild, M., *Astr. J.*, **139**, 587 (1964).

¹⁰ Shapiro, I. I., Smith, W. B., Ash, M. B., Ingalls, R. P., and Petten-gill, G. H., *Phys. Rev. Lett.*, **26**, 27 (1971).

Origin of the Alps and Western Mediterranean

K. J. HSÜ

Geological Institute, ETH, Zürich

The origin of the Alpine geosyncline may have been related to the eastward motion of Africa relative to Europe during the Mesozoic, and orogenesis in western Europe could have been caused by a westward and then a northward movement of Africa.

In contrast to mountains such as the Himalayas and the Andes, the Alpine chain snakes its way through Europe with numerous bends, convergences and divergences^{1,2}. Compared with the Atlantic, the opposite shores of which fit together almost perfectly, the Mediterranean is bounded by three peninsulas on the north, which run at right angles to the trend of the sea. To make matters worse the Mediterranean Ridge has proved to be a compressional feature³, quite distinct from the normally extensional mid-ocean ridges.

The complex state of the Mediterranean reflects its past history and conventional interpretations have related the origin of the Alps and the Mediterranean to the interaction between the European and African plates. In many of these

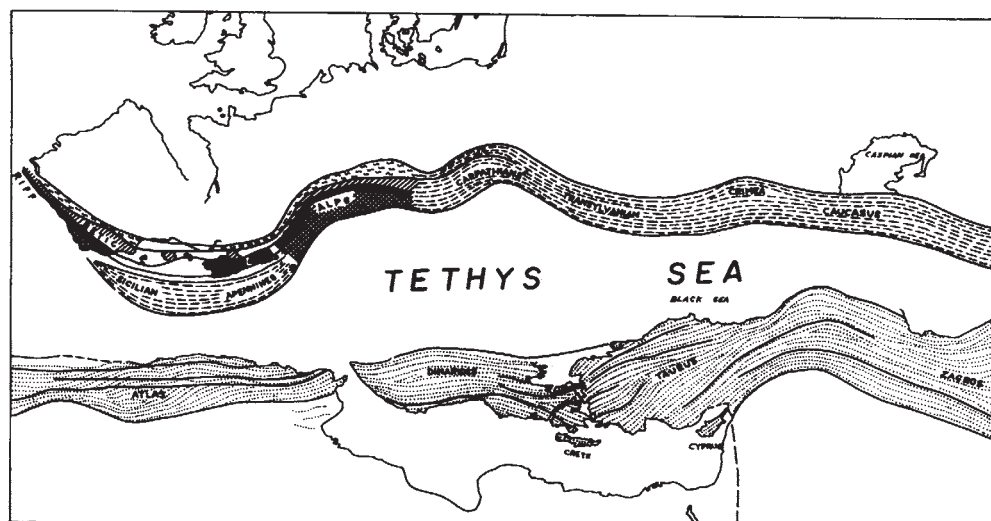
interpretations a broad, east-west sea (the Tethys Sea) was placed between the two continents (Fig. 1), and the Alpine folding was postulated as a consequence of the northward march of Africa^{4,5}. Such a broad new picture left many of the intricacies unexplained and the important geological fact that a deep oceanic seaway did not exist in western Europe (the Alps, Apennines, and Betic Cordilleras) before the Jurassic was ignored.

Geometrical Analysis

Staub⁴ considered the winding Alpine trend as a trait inherited from the European Hercynian system and thereby evaded the issue. Argand⁶, however, presented a brilliant geometrical analysis and invoked continental drift and rotation of microcontinents to account for the puzzling geography of the western Mediterranean. But his kinematical analyses were based on geological data alone and have been proved untenable by recent geophysical and oceanographical discoveries. This article is an attempt to modify Argand's scheme; if the drifting chronology is revised, the geometrical displacements postulated by Argand acquire a different geological significance.

Argand⁶ postulated that Africa was pushed over Eurasia in the early Tertiary, and formed the Alpine and Carpathian mountains while Eurasia was still attached to North America. After the climax of the Alpine deformation in the Oligocene,

Fig. 1 A current interpretation of the Mesozoic palaeogeography of the Tethyan geosyncline (after ref. 5).



the northern margin of the African plate was bounded by the ophiolite zone of the Rif, the Betic Cordilleras, Corsica, Liguria, and the Pennine Alps in the west, and by the Carpathian Flysch in the north-east (Fig. 2). The Mediterranean was supposedly reduced into a network of epicontinental seas and Italy, Corsica, Sardinia, the Balearic Islands, and Spain were all crowded to one corner in South-Western Europe and joined to France and the stable Europe. He then postulated the start of an important phase of extension during the Late Oligocene or Early Miocene, when Africa moved south-eastward relative to Europe, inducing the rift of the Alboran Sea, causing the rotation of Spain, and creating thereby the Bay of Biscay. Meanwhile Corso-Sardinia and Italy somehow drifted counter-clockwise and left behind the Balearic and Tyrrhenian Basins. The collision of the Italian Peninsula and the Balkan block also accounted for the genesis of the Apennines.

Sphenochasm

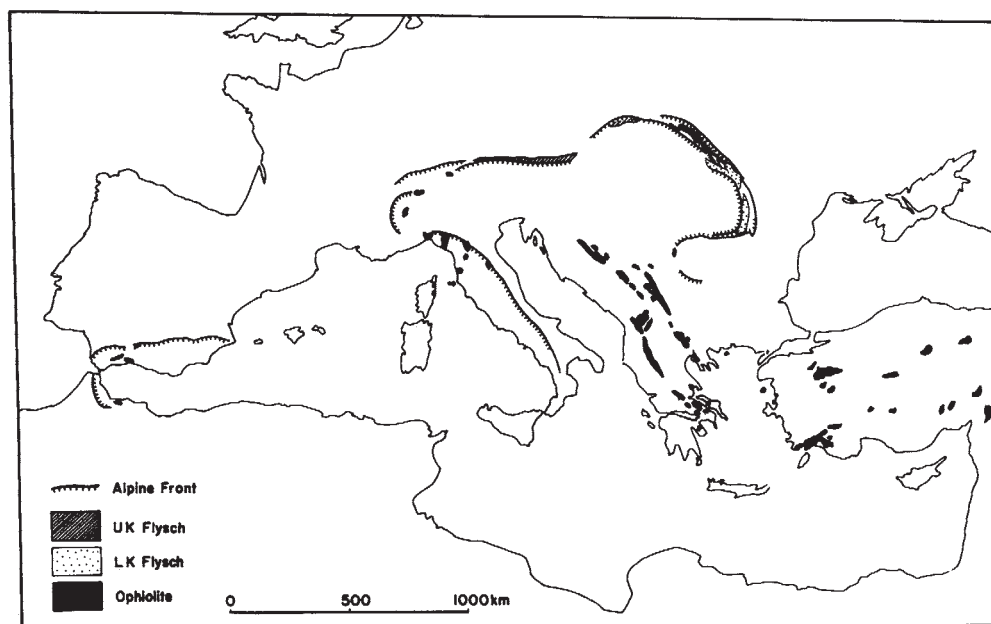
Argand's scheme was slightly modified by Carey⁵, who spoke of a Ligurian sphenochasm. The closing of such a sphenochasm should place the Corso-Sardinia Block next to the Gulf of Lyon but Carey put this block east rather than north of the Balearic Island. His reconstruction would make it almost impossible to

correlate the Subbetic Zone of the Balearic Islands with the Pennine Zone of Corsica and Elba. Carey was also obliged to postulate an unwieldy set of shear motions, which could not easily be related to a rational driving mechanism. Such a scheme seems to me to be a backward step and does not represent any improvement.

But Argand's geometrical postulates do have considerable merit. The idea that the genesis of the Bay of Biscay was related to the rotation of Spain can be considered proved⁷. Several lines of evidence have supported the notion of the counter-clockwise rotations of Corso-Sardinia^{8,9} and of Italy¹⁰. Most of Argand's chronological postulates have, however, been contradicted by new evidence; the opening of the Bay of Biscay, for example, could not have started as late as the Oligocene¹¹. All the evidence indicates a Mesozoic date probably during the Late Jurassic and certainly not later than the Early Cretaceous^{12,13}. There is likewise no compelling reason for assuming a Middle Tertiary rotation of Italy away from Corso-Sardinia.

The rotations postulated by Argand were supposedly caused by a displacement of the African plate south-eastwards relative to stable Europe, yet such a relative motion could not have taken place during the Late Tertiary when Europe was moving away from Greenland at such a rate that a north-south compression

Fig. 2 Distribution of Mesozoic ophiolites and Flysch in Europe.



between Europe and Africa resulted¹⁴. (Pittman's important work is being completed in collaboration with Ryan, Dewey, and Bonin. Their kinematical analysis of the North Atlantic opening defined the succession of relative movements between Africa and Europe and thus paved the way for my attempt to revise Argand's chronology on the genesis of the western Mediterranean.) In fact this motion still seems to be continuing today³.

Relative Motion

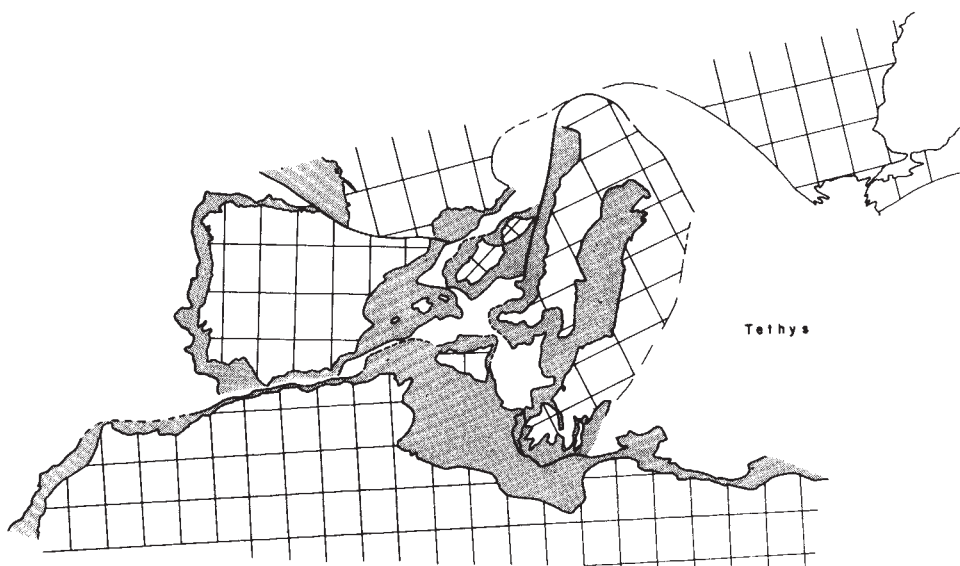
The magnetic anomaly patterns of the North Atlantic and the results of the JOIDES drilling there indicate that the history of the relative motion between Africa and Europe can be subdivided into three stages. These are (1) the Middle Jurassic to the Middle Cretaceous (165 to 80 m.y.) when the Middle Atlantic was being opened as Africa pulled away from North America^{3,16}. The motion of Africa relative to Europe was eastward, because Europe was still joined to North America. (2) The Middle Cretaceous to the Late Eocene (80 to 40 m.y.) when Greenland and Europe were also being separated from North America¹¹. The motion of Africa relative to Europe

One of the most significant drilling results of the JOIDES Legs X and XIII was the discovery of a Lower Cretaceous and Upper Jurassic pelagic sequence above an ophiolitic oceanic basement, almost identical to the coeval Alpine sequences of western Europe^{3,19}. The eastward march of Africa which rotated the various microcontinents also permitted the opening of the Atlantic. The geological history of the Atlantic and the Alpine regions ran a parallel course as long as this movement continued, but the early Atlantic could not have been very broad, nor was the circulation quite normal as suggested by the abundant pyritized fossils and other Red Sea type of metallic deposits in the JOIDES cores^{3,19}. Slump blocks from the nearby steep walls have also been observed³. But the Mesozoic Alpine geosynclines—those inland seas, a few hundred kilometres wide, which might be considered as the ancestral western Mediterranean (Fig. 4)—did not reach oceanic proportions¹⁵.

Different Ophiolites

Alpine geologists have long recognized two different sets of ophiolites in the European Tethyan system⁴ (Fig. 2). The

Fig. 3 Early Jurassic palaeogeography of Europe before the start of the motion of Africa eastward relative to stable Europe.



was then westward, as Europe was separating from North America at a greater rate¹⁴. (3) The Late Eocene to the Present (40 m.y. to now) when Europe was separating from Greenland. The motion of Africa relative to Europe was east and then northward.

This kinematical analysis agrees well with the geological information. Stratigraphical evidence suggests unequivocally that western Europe was covered by a shelf-sea during the Triassic, when there was no oceanic seaway separating Spain from Corso-Sardinia, or Italy from stable Europe¹⁵. All those future microcontinents were still tucked in between Europe and Africa west of the open Tethyan Sea (Fig. 3). Thus the stage was set for Argand's postulated extensional phase in the Early Jurassic, not in the Late Oligocene¹⁶. Evidence of the start of crustal extension in the Jurassic is given by the first deposition of continental margin sediments in the Alps¹⁵. By the Middle Jurassic the oldest ophiolites of western Europe were emplaced and the microcontinents were split up, thus creating the Alpine, Apennine, Betic, and Pyrenees geosynclines (Fig. 4). The ophiolites in those deep-sea troughs were overlain by Upper Jurassic and Lower Cretaceous radiolarian and nannofossil oozes^{15,17}, forming the typical sequences of the Steinmann's Trinity¹⁸.

Penninic ophiolites of the Subbetic Cordilleras, the Apennines, and the Alps could not be traced beyond the central Carpathians but a second set of Bosnian ophiolites marks the approximate position of a suture zone in the Balkans. Whereas the Penninic ophiolites are not older than Jurassic, the Bosnian ophiolites could be Triassic or even older, and Palaeozoic ophiolites from Anatolia are known²⁰. These palaeozoic ophiolites were obviously the remnants of the once extensive Tethyan Sea between Eurasia and Africa. As Africa pushed the Greco-Italian Block eastward and caused it to rotate counterclockwise, the Tethyan Ocean north-east of the Block was consumed as is evidenced by the Jurassic and Lower Cretaceous Flysch of the eastern Carpathians^{21,22} (Fig. 2).

The Greco-Italian microcontinent was joined onto the African plate during the Middle Cretaceous, and has moved with it ever since. (In ref. 23, D. P. McKenzie did not define precisely the present plate boundary between Africa and Europe. The possibility that Italy and western Balkan are a part of the African plate was discussed with him after the Royal Society Symposium on Plate Tectonics in February 1971. He found that my idea is not contradicted by available data on the seismicity of these regions.) Geological evidence also suggests a Late Cretaceous westward motion of the Italo-African plate;

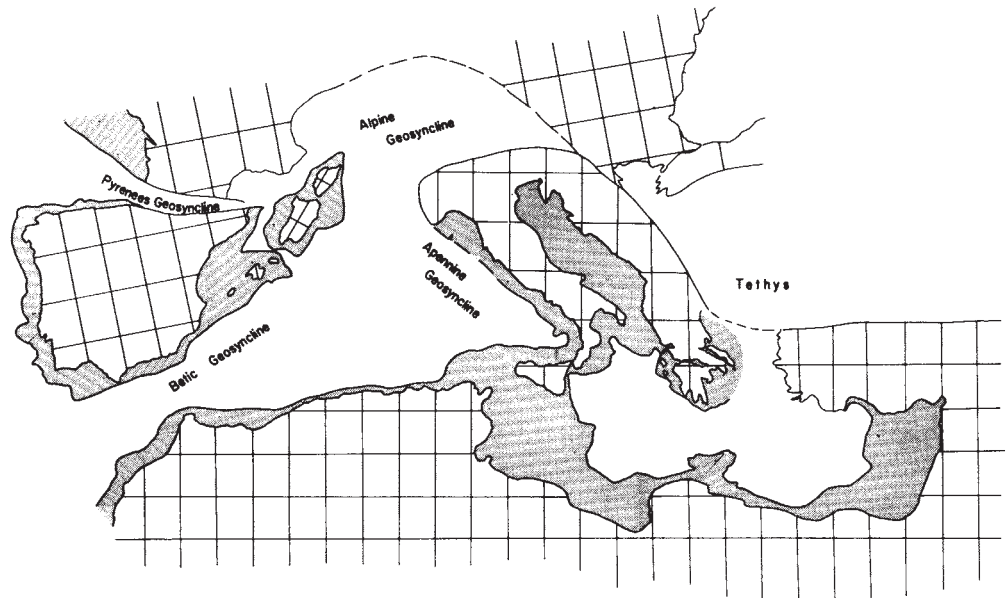
the Austroalpine nappes of the Austrian and Swiss Alps represented the original northern margin of this Cretaceous Italo-African plate and these rigid basement blocks were thrust westward onto the Pennine ophiolites during orogenic movements which started during the Middle Cretaceous and culminated in the Late Eocene²⁴. The Pennine ophiolites were almost completely consumed by the beginning of the Oligocene, at which time this northern margin of the Italo-Africa plate had a sutured contact with Europe; the debris eroded from the Austroalpine nappes were then being deposited in nearly Molasse troughs on the southern margin of the European plate.

The western margin of the Italo-African plate overrode the ophiolites of the Apennine geosyncline, giving rise to the Late

Africa relative to Europe¹⁵. Geological evidence of this movement is clear in the well investigated Alps, Apennines and Carpathians, but the geology is less well known elsewhere. The northward motion of Africa has continued to this day and has resulted in the underthrusting of the Hellenic Trench^{3,23} and the elevation of the Mediterranean Ridge³. We are, however, baffled by the latest history of the Alboran Basin, where geological evidence seemed to indicate a late Cenozoic rifting^{3,6} in contrast to the present state of compression²³.

We are presenting a very speculative idea in an attempt to rationalize some of the very puzzling aspects of the Mediterranean geography and geology. We are in general agreement with some of the geometrical postulates of Argand, but disagree

Fig. 4 Middle Cretaceous palaeogeography of Europe when the westward motion of Africa relative to stable Europe began.



Eocene deformation of the ophiolitic mélanges of Liguria and on the islands of Elba and Corsica²⁵⁻²⁷. The first deformation of the Betic must also have taken place during this westward motion of Africa, when the Betic ophiolitic geosyncline was almost eliminated. Unfortunately, geological evidence from this relatively poorly investigated orogenic belt is meagre.

Westward Movement

When the Iberian Block was joined to the African Plate during the Eocene and began to share its westward movement, the deformation of the Pyrenees commenced and may have continued during the subsequent northward movement of Africa. The present plate boundary between Europe and Africa seems to be located near the Strait of Gibraltar²³ and may represent a later development.

The kinematics of the Corso-Sardinia motion is very uncertain. Sedimentological evidence indicated that this block was still partly connected to southern France during the Early Oligocene⁹, and palaeomagnetic evidence suggested a post-Oligocene rotation of Sardinia²⁵. The JOIDES drilling results are also not inconsistent with a young age for the northern Balearic Basin³. Furthermore no Alpine ophiolite zone exists between this block and the stable Europe. Thus we tended to agree with Argand and modern Italian geologists²⁵ that the Corso-Sardinia block was not separated until the culmination of ophiolite deformation in the Late Eocene.

Our understanding of the origin of the Alps and the Mediterranean has been made difficult because of the last phases of the Alpine deformation which involved a northward motion of

with him on the timing of events. We postulate that the origin of the Alpine geosyncline was related to the Mesozoic eastward motion of Africa relative to Europe, and that the orogenic movements in western Europe were caused by a westward and later northward motion of Africa relative to Europe. This hypothesis is not offered as a final solution; the proposal of some definite ideas will, however, focus attention on several areas where geological and geophysical data could be synthesized and freshly gathered to check, improve, or disprove this working hypothesis.

Received March 19, 1971.

- ¹ Suess, E., *Das Antlitz der Erde* (Freytag, Leipzig, 1885).
- ² Bucher, W., *The Deformation of the Earth's Crust* (Princeton University Press, 1933).
- ³ Ryan, W. B. F., Hsü, K. J., Nesteroff, V., Pautot, G., Wezel, F., Lort, J., Cita, M., Maync, W., Stradner, H., and Dumitrica, P., *Geotimes*, **15**, 12 (December 1970).
- ⁴ Staub, R., *Der Bewegungsmechanismus der Erde* (Bornträger, Berlin, 1928).
- ⁵ Carey, S. W., *Symp. Continental Drift*, 177 (Tasmania, 1958).
- ⁶ Argand, E., *Cong. Geol. Intern., XIIIe, Belgique, C.R.*, 171 (1924).
- ⁷ *Proc. Symp. sur l'Histoire Structurale du Golfe de Gascogne, December 1970* (Rueil-Malmaison, France) (in the press).
- ⁸ Zijdeveld, J. D. A., De Jong, K. A., and van der Voo, R., *Nature*, **226**, 933 (1970).
- ⁹ Stanley, D., and Mutti, E., *Nature*, **218**, 6 (1968).
- ¹⁰ Zijdeveld, J. D. A., and De Jong, K. A., *Geologie en Mijnbouw*, **48**, 559 (1969).
- ¹¹ Laughton, A. S., Berggren, W. A., Bensen, R., Davis, T. A., Franz, U., Musich, L., Perch-Nielsen, K., Ruffman, A., van Hinte, J. E., and Whitmarch, R. B., *Geotimes*, **17**, 10 (November 1970).

- ¹² Zijdeveld, J. D. A., and De Jong, K. A., *Symp. Golfe de Gascogne*, (Rueil-Malmaison, 1970).
- ¹³ Stauffer, K. W., and Tarling, D. J., *Symp. Golfe de Gascogne*, (Rueil-Malmaison, 1970).
- ¹⁴ Pittman, W., *Geol. Soc. Amer. Ann. Meetings Program*, 2, 752 (1970).
- ¹⁵ Trümpy, R., *Geol. Soc. Amer. Bull.*, **71**, 843 (1960).
- ¹⁶ Smith, A. G., *Proc. Roy. Soc. Symp. Plate Tectonics*.
- ¹⁷ Maxwell, J. C., in *The Megatectonics of Continents and Oceans* (edit. by Johnson, H., and Smith, B. L.) (Rutgers University Press, 1970).
- ¹⁸ Bailey, E. B., and McCallien, W. J., *Trans. Roy. Soc. Edinburgh*, **65**, 315 (1963).
- ¹⁹ Ewing, J., Hollister, C., Hathaway, J., Paulus, F., Lancelot, Y., Habib, D., Poag, C. W., Luterbach, H. P., Worstell, P., and Wilcoxon, J. A., *Geotimes*, **15**, 14 (1970).
- ²⁰ Arni, P., *Z. Inst. Lagerstättenforschung Türkei*, **3**, 472 (1942).
- ²¹ Birkenmajer, K., *Bull. Acad. Polonaise Sci.*, **4**, 451 (1957).
- ²² De Raaf, J. F. M., *The Science of Petroleum*, **6**, 10 (Oxford University Press, 1953).
- ²³ McKenzie, D. P., *Nature*, **226**, 239 (1970).
- ²⁴ Trümpy, R., Manuscript in press for van Bemmelen Festschrift volume.
- ²⁵ Boccaletti, M., and Guazzone, G., *Mem. Soc. Geol. Italiana*, **9**, 177 (1970).
- ²⁶ Hsü, K. J., *Bull. Amer. Assoc. Petrol. Geol.*, **51**, 65 (1967).
- ²⁷ Sestini, G., *Sediment. Geol.*, **4**, 203 (1970).

Molecular Mechanisms of Chromosome Pairing, Folding and Function

DAVID E. COMINGS & ARTHUR D. RIGGS

Departments of Medical Genetics and Biology, City of Hope National Medical Center, Duarte, California 91010

Allosteric proteins are proposed which, on binding to specific DNA sites, undergo conformational changes that enable them to bind to each other and to control pairing, condensation, and genetic function of chromosomes.

IN this article we wish to show that allosteric DNA binding regulatory proteins such as the *lac* repressor are admirably suited to direct eukaryotic chromosome folding and homologous pairing in addition to controlling gene expression. The significant properties include tight and specific binding to DNA at a unique base sequence, tight and specific interaction between their subunits, their sensitivity to inducers or effector ligands and the conformational changes induced by DNA binding. Thus these proteins would bind to specific DNA sites and undergo allosteric changes that would enable them to bind to each other or to change enzymatic activity. They could be called "nucleosteric proteins". On this basis, it is possible to propose mechanisms to account for a number of biological phenomena, such as homologous chromosome pairing, that have previously been extremely difficult to explain.

Pre-meiotic and Somatic Pairing

Chromosome pairing during meiosis in eukaryotes probably occurs in three stages^{1,2}. These are pre-meiotic pairing or alignment of homologous chromosomes, chromosome synapsis mediated by the synaptonemal complex, and intimate molecular pairing between short stretches of DNA containing similar base sequences. The last stage may be initiated by the formation of heteroduplex DNA molecules^{3,4} and may lead to recombination. The synaptonemal complex seems to enhance the efficiency of molecular pairing, because recombination associated with meiotic pairing is much more frequent than that associated with somatic pairing^{5,6}. Yet chromosome synapsis mediated by the synaptonemal complex seems to be non-specific, for it can also pull together non-homologous chromo-

somes^{1,2,7,8}. It seems to act by zippering together chromosomes which have their homologous telomeres attached next to each other at the nuclear membrane^{1,9,10}. Thus one of the most important aspects of meiotic chromosome pairing concerns the mechanism by which homologous chromosomes become partially aligned before the formation of the synaptonemal complex¹¹⁻¹⁷. This may be the same mechanism that is responsible for the precise somatic pairing that is seen clearly in diptera species. It must be capable of acting throughout the length of the chromosome, since complex inversions in *Drosophila* show perfect homologous somatic pairing in spite of the absence of the synaptonemal complex.

Fig. 1 illustrates a scheme by which this chromosomal pairing might be accomplished. It requires no new phenomena or forces, only the three properties already mentioned. We suggest that genes exist which code for "pairing proteins", each capable of binding specifically to a unique base sequence. On binding to DNA, these protein pairs undergo allosteric changes in surface configuration which enhance their complementarity. Their subsequent association enables homologous pairing between specific segments of the chromosomes. The allosteric modification after DNA binding prevents association of the unbound pairing proteins and ensures that the only proteins in the nucleus that can associate are those that are bound to the two homologous DNA sites. It is suggested that a series of these DNA sites and their specific pairing proteins occur throughout each chromosome. The intriguing aspect of this model is that it provides a mechanism for the pairing of native DNA molecules that is as specific and thermodynamically stable as the Watson-Crick pairing of half DNA helices.

The alternative mechanisms of homologous pairing are less appealing. For example, it could be assumed that this type of chromosome pairing is mediated by the presence of specific nucleases that keep homologous regions of DNA partially unpaired and that pairing would be mediated by heteroduplex formation between these single stranded whiskers. The difficulty with this mechanism is that it would lead to a high frequency of recombination when in fact recombination between somatically paired chromosomes is rare. Furthermore, if pre-meiotic pairing is accomplished in this manner, why should a synaptonemal complex evolve to facilitate essentially the same thing

at a later stage in meiosis? A second alternative could involve the formation of a bivalent pairing protein with two identical DNA recognition sites. A fusion gene formed by unequal crossing over could produce such a protein. The bivalent pairing protein would attach first to one homologous site, then to the other. The difficulty with this scheme is that the efficiency of such a pairing protein would be reduced by competition between partially bound and unbound molecules. Furthermore, it seems unlikely that a whole series of such genes could be produced by the hazardous route of unequal crossing over.

Chromosome Folding

The proposed mechanism of chromosome pairing involves recognition sites that exist as single copies per haploid set, but it is equally plausible to propose a repetitious sequence as the recognition site, thus providing a mechanism for chromosome folding or condensation (Fig. 2). A DNA sequence repeated many times throughout the genome could recognize a specific folding protein synthesized before mitosis. This protein would

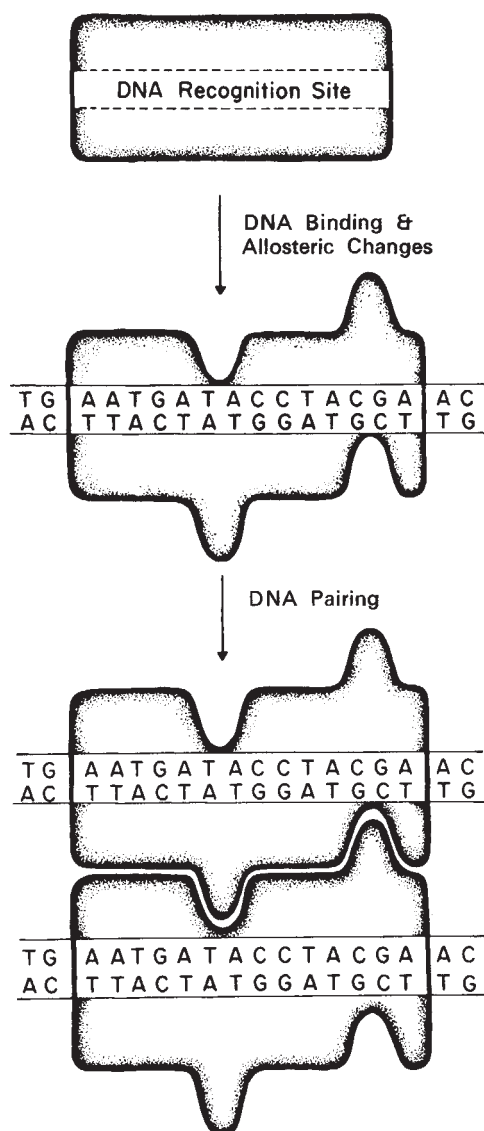


Fig. 1 A proposed mechanism of homologous chromosome pairing in somatic and premeiotic cells. Pairing proteins are synthesized which possess a DNA recognition site capable of firm binding to a specific base sequence. After binding, allosteric changes occur which allow the protein to bind itself, thus initiating the homologous pairing of that site. It is suggested that a series of such regions with their specific pairing proteins are located throughout the length of each chromosome.

be capable of self-association only after binding DNA. The net-like arrangement of chromatin that would be formed would closely resemble that seen in whole mount electron microscopy of mitotic chromosomes¹⁸⁻²⁰. Such a mechanism could provide an explanation for the condensation of interphase chromatin when the nucleus is fused with a cell in mitosis²¹. The aggregation could be destroyed at the end of mitosis by the synthesis of new proteins that interfere with the binding of the folding protein to DNA, releasing them for destruction before the next mitosis. Shifts in cation concentration or low molecular weight ligands could also accomplish this release. Similar mechanisms could be evoked to account for the binding and release of chromatin to the nuclear membrane during the cell cycle²²⁻²⁴. The attachment of the chromosomes to localized portions of the nuclear membrane would help to ensure that they condensed on themselves rather than fusing to each other.

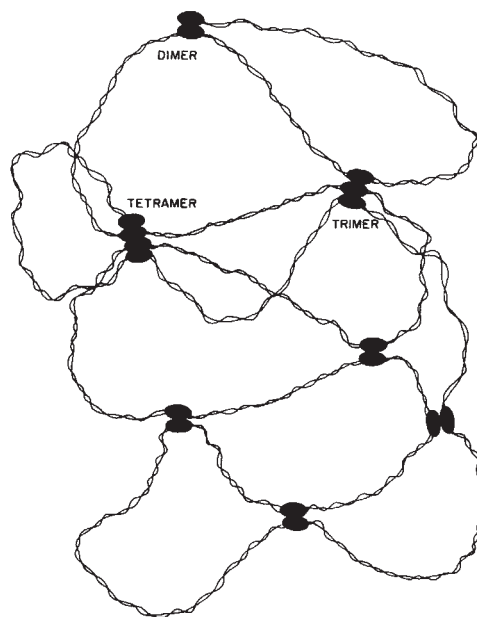


Fig. 2 Chromosome condensation or folding during mitosis could be accomplished by a repetitious DNA sequence that is recognized by a "folding protein". After binding to DNA this protein would undergo allosteric changes that allow it to bind to itself (see text).

There may be at least two categories of folding proteins, heterochromatin specific and euchromatin specific. Heterochromatin is chromatin that remains condensed and densely staining throughout interphase²⁵. Many lines of evidence suggest that specific types of DNA occur within heterochromatin²⁶⁻²⁹. Thus, it is reasonable to propose that it may contain specific DNA recognition sites not found in euchromatin and that the folding proteins that bind to these sites are synthesized throughout the cell cycle. This could provide a simple mechanism for keeping this chromatin permanently condensed. It is also consistent with the observation that heterochromatin may be condensed less during mitosis than euchromatin³⁰ suggesting that its folding sites have already been taken up and that those utilized for the rest of the chromosome are different and more efficient.

Properties of Allosteric Proteins

We have proposed mechanisms that involve proteins with the properties of binding and specificity that are, as we shall show here, typical of repressor proteins. It has now been clearly shown that regulatory proteins such as the *E. coli lac* repressor can bind tightly and specifically to a unique base sequence in

double stranded DNA³¹⁻³⁴ and similar results have been obtained for the λ and 434 phage repressor³⁵⁻³⁷. The *lac* repressor can identify the sequence of perhaps twenty base pairs corresponding to the *lac* operator. Twenty base pairs would be more than adequate to provide unique sequences even in genomes 1,000 times as large as *E. coli*. The *lac* repressor binds very tightly to the *lac* operator. At low salt concentrations, the dissociation constant can be less than 10^{-13} M, corresponding to a free energy change of 18 kcalories. In aqueous solution many covalent bonds are not nearly as strong. At higher salt concentrations the binding is weakened but remains very tight; in 0.2 M KCl the dissociation constant is still 10^{-11} M³². The binding of sequence specific proteins to DNA is certainly strong enough to hold homologous chromosomes in register, especially if there is cooperative binding at many distinct sites.

The bonding forces involved in the association of identical protein subunits to form oligomers are also sufficiently specific and strong. That the specificity and speed of association is adequate is made clear by the fact that even though there are only about ten copies of tetrameric repressor per haploid *E. coli* cell, and the concentration of subunits (which bind IPTG and would be detected) is even less^{38,39} the repressor subunits assemble quickly to form active tetramers⁴⁰. The presence of all other *E. coli* proteins does not interfere. The tightness with which repressor subunits bind to each other has not been ascertained, but since assembly takes place in such dilute solutions the bonding forces must be strong. The assembly of β -galactosidase is also rapid in dilute solution⁴¹. For other proteins, the binding between subunits can be as high as 13 kcalories/mol of dimer formed⁴².

Allosteric proteins are by definition sensitive to effector ligands⁴³. These ligands can be substrates or compounds unrelated to the substrate. Effector ligands cause or stabilize allosteric transitions which result in a changed conformation of the polypeptide chains^{44,45}. Very often these conformational changes affect the subunit interaction and can promote either association or dissociation^{43,44,46}. The *lac* repressor is clearly an allosteric protein in that it has a galactoside binding site distinct from the DNA binding site³³. Furthermore, these effector ligands can change considerably the stability of repressor-operator complex. One inducing ligand, IPTG, is known to increase more than 100 times the rate of repressor-operator dissociation, while another, ONPF, stabilizes the complex³³.

There is only limited evidence at present that the conformation of the operator-bound repressor is different from that of the free repressor, but it has been shown that the apparent affinity of IPTG for DNA bound repressor is considerably less than for repressor free in solution. This strongly suggests differences in conformational state (A. D. R., unpublished data). That DNA should act as an allosteric effector, changing the conformation of the bound protein, is very plausible. Substrates often behave as allosteric effectors and when such strong binding forces are involved, it would, we feel, be more surprising if changes were not observed.

Chromosome Function

The above postulate that DNA may act as an allosteric effector leads us to suggest a novel mechanism that can account for positive control of gene activation: there may be enzymes which alter histones and which are active only when associated with their DNA binding site; these would be known, therefore, as DNA-dependent histone inactivating enzymes. The alteration of histone binding could be produced by a number of reactions including methylation, acetylation and phosphorylation^{47,48}. In the micro-environment around the specific DNA binding site, the concentration of histone inactivating enzyme would be very high and the non-specific repression by histones would be greatly reduced, permitting gene expression. DNA would then serve the purpose of both activating and fixing the spatial location of regulatory enzymes.

In conclusion, we propose the existence of a class of nucleosteric proteins which bind to specific DNA sites and, on binding, undergo allosteric changes that allow them to bind to each other or to change enzymatic activity. These proteins could play a role in initiating homologous chromosome pairing, provide for the condensation of chromosomes at mitosis, condensation of heterochromatin during interphase, and they could be involved in both positive and negative control of gene function. Techniques are available to test for the existence of proteins with these unique properties.

This work was supported by grants from the National Institutes of Health.

Received April 26, 1971.

- ¹ Comings, D. E., and Okada, T. A., *Nature*, **227**, 451 (1970).
- ² Comings, D. E., and Okada, T. A., in *Adv. Cell Mol. Biology* (edit. by DuPraw, E. J.) (Academic Press, New York, in the press).
- ³ Whitehouse, H. L. K., and Hastings, P. J., *Genet. Res.*, **61**, 27 (1965).
- ⁴ Holliday, R., in *Replication and Recombination in Genetic Material* (edit. by Peacock, W. J., and Brock, R. D.), p. 157 (Australian Acad. Sci., Canberra, 1968).
- ⁵ Pontecorvo, G., *Trends in Genetic Analysis* (Oxford University Press, 1959).
- ⁶ Stern, C., *Genetics*, **21**, 625 (1936).
- ⁷ Menzel, M. Y., and Price, J. M., *Amer. J. Bot.*, **53**, 1079 (1966).
- ⁸ Ting, Y. C., *Genetics*, **61**, s58 (1969).
- ⁹ Comings, D. E., and Okada, T. A., *Chromosoma*, **30**, 269 (1970).
- ¹⁰ Moens, P. B., *Chromosoma*, **28**, 1 (1969).
- ¹¹ Smith, S. G., *Canad. J. Res.*, **20**, 221 (SD) (1942).
- ¹² Moens, P. B., *Chromosoma*, **15**, 231 (1964).
- ¹³ Feldman, M., *Proc. U.S. Nat. Acad. Sci.*, **55**, 1447 (1966).
- ¹⁴ Maguire, M., *Chromosoma*, **21**, 221 (1967).
- ¹⁵ Brown, W. V., and Stack, S. M., *Bull. Torrey Bot. Club.*, **95**, 369 (1968).
- ¹⁶ Driscoll, C. J., and Darvey, N. L., *Science*, **169**, 290 (1970).
- ¹⁷ Lu, B. C., and Raju, N. B., *Chromosoma*, **29**, 305 (1970).
- ¹⁸ DuPraw, E. J., *Nature*, **206**, 338 (1965).
- ¹⁹ Comings, D. E., and Okada, T. A., *Cytogenetics*, **9**, 436 (1970).
- ²⁰ Comings, D. E., and Okada, T. A., *Cytogenetics*, **9**, 450 (1970).
- ²¹ Johnson, R. T., and Rao, P. N., *Nature*, **226**, 717 (1970).
- ²² Comings, D. E., *Amer. J. Human Genet.*, **20**, 440 (1968).
- ²³ Comings, D. E., and Okada, T. A., *Exp. Cell Res.*, **62**, 293 (1970).
- ²⁴ Comings, D. E., and Okada, T. A., *Exp. Cell Res.*, **63**, 62 (1970).
- ²⁵ Heitz, E., *Jahrb. F. Wissensch. Bot.*, **69**, 762 (1928).
- ²⁶ Yasmineh, W. G., and Yunis, J. J., *Exp. Cell Res.*, **59**, 69 (1970).
- ²⁷ Pardue, M. L., and Gall, J. G., *Science*, **168**, 1356 (1970).
- ²⁸ Jones, K. W., *Nature*, **225**, 912 (1970).
- ²⁹ Mattoccia, E., and Comings, D. E., *Nature New Biology*, **229**, 175 (1971).
- ³⁰ Comings, D. E., *Exp. Cell Res.*, (in the press).
- ³¹ Gilbert, W., and Muller-Hill, B., *Proc. U.S. Nat. Acad. Sci.*, **56**, 1891 (1966).
- ³² Riggs, A. D., Suzuki, H., and Bourgeois, S., *J. Mol. Biol.*, **48**, 67 (1970).
- ³³ Riggs, A. D., Newby, R. F., and Bourgeois, S., *J. Mol. Biol.*, **51**, 303 (1970).
- ³⁴ Riggs, A. D., Bourgeois, S., and Cohn, M., *J. Mol. Biol.*, **53**, 401 (1970).
- ³⁵ Ptashne, M., *Proc. U.S. Nat. Acad. Sci.*, **57**, 306 (1967).
- ³⁶ Ptashne, M., *Nature*, **214**, 232 (1967).
- ³⁷ Pirrotta, B., and Ptashne, M., *Nature*, **222**, 541 (1969).
- ³⁸ Muller-Hill, B., Crapo, L., and Gilbert, W., *Proc. U.S. Nat. Acad. Sci.*, **59**, 1259 (1968).
- ³⁹ Jobe, A., Riggs, A. D., and Bourgeois, S., *J. Mol. Biol.* (in the press).
- ⁴⁰ Branscombe, E. W., and Stuart, R. N., *Biochem. Biophys. Res. Commun.*, **32**, 731 (1968).
- ⁴¹ Adamson, L., Gross, C., and Novick, A., in *The Lactose Operon* (edit. by Beckwith, J. R., and Zipser, D.), 317 (Cold Spring Harbor Laboratory, New York, 1970).
- ⁴² Klotz, I. M., Langerman, N. R., and Darnall, D. W., *Ann. Rev. Biochem.*, **39**, 25 (1970).
- ⁴³ Monod, J., Changeux, J. P., and Jacob, F., *J. Mol. Biol.*, **6**, 306 (1963).
- ⁴⁴ Koshland, D. E., and Neet, K. E., *Ann. Rev. Biochem.*, **37**, 359 (1968).
- ⁴⁵ Perutz, M. F., *Nature*, **228**, 726 (1970).
- ⁴⁶ Tao, M., Salas, M. L., and Lipmann, F., *Proc. U.S. Nat. Acad. Sci.*, **67**, 408 (1970).
- ⁴⁷ Allfrey, V. G., Faulkner, R., and Mirsky, A. E., *Proc. U.S. Nat. Acad. Sci.*, **51**, 786 (1964).
- ⁴⁸ Kleinsmith, L. J., Allfrey, V. G., and Mirsky, A. E., *Proc. U.S. Nat. Acad. Sci.*, **55**, 1182 (1966).

LETTERS TO NATURE

PHYSICAL SCIENCES

Discrimination of Earthquakes and Explosions using Short Period Seismic Array Data

THE problem of identification of earthquakes and underground explosions on a global basis, for seismic events with body wave magnitudes (m_b) below 4.7 has not been solved^{1,2}, although many partially successful discriminants have been proposed. A positive identification parameter based solely on the short period seismic data is interesting as the equivalent long period energy detection threshold is often difficult to achieve, particularly at teleseismic distances.

Using data recorded at the Yellowknife short period seismic array in northern Canada, a simple two dimensional identification parameter has been found that completely separates a population of underground explosions ($4.2 \leq m_b \leq 6.2$) at the Urals, Caspian, East Kazak, Sinkiang and Novaya Zemlya test sites from a population of shallow ($h \leq 50$ km) earthquakes ($4.3 \leq m_b \leq 6.0$) within about 15° of these test sites. The data were examined in both the time and frequency domain for optimum identification parameters that could be used either singly or in a multivariate discriminant. Short period multivariate discriminants have been published in various reports of the Lincoln Laboratory of MIT³ and have provided positive identification for events of m_b 4.7.

For the Yellowknife data, the optimum discriminant in the time domain was found to be a complexity parameter C defined as

$$C = \frac{\sum_{t=2}^{t=35} |x(t)|}{\sum_{t=0}^{t=2} |x(t)|}$$

where $|x(t)|$ is the rectified sum of the nineteen signal channels of the array, phased to that beam that maximizes the correlogram of the P phase⁴. The algorithm, which is used to define $t=0$, is based on the correlogram and $t=0$ is assumed where the value of the correlogram first rises above a given bias value. Using this automated procedure for defining $t=0$, results in a starting time that is always fractionally later than the actual onset time of the P phase, the difference is of the order of 0.2 s. This complexity parameter is the same as one used at the Lincoln Laboratory⁵ except for the definition of $t=0$.

Using the same data set, the optimum discriminant in the frequency domain was found to be the third moment of frequency (TMF), defined as

$$\text{TMF} = \sum_{n=1}^{n=19}$$

$$\left(3 \sqrt{\sum_{f=0.32}^{f=5.0} (A_n(f) - B_n(f)) \cdot f^3} / \sum_{f=0.32}^{f=5.0} (A_n(f) - B(f)) \right) / 19$$

where $A_n(f)$ is the spectral amplitude at frequency f (Hz) of the first 128 samples of the signal (twenty samples per second), of the data channel n , and $B_n(f)$ is the mean spectral amplitude,

at the same frequency and on the same data channel, of two noise samples taken in the 20 s before the onset of the event. If $A_n(f) < B_n(f)$, $A_n(f) - B_n(f)$ is set to zero. The spectral amplitudes were obtained using the Cooley-Tukey algorithm, the data windows are cosine tapered by 12.5% at each end. Weichert⁶, without using the $B_n(f)$ term has also found the third moment of frequency to be the optimum discriminant in the frequency domain.

The earthquake population of fifty-one shallow events was selected from the 1969 USCGS PDE lists using all those events available from the Yellowknife array digital tape library. The explosion population of twenty-seven covers the years 1968 to 1970 and likewise was obtained from all Eurasian events on the digital tape library. For some events, not all of the nineteen array channels were operational or free from system noise, but these events were still retained in the populations in order to give a realistic data sample; however, inoperative or excessively noisy array channels were omitted from the calculations of C and TMF.

Using the complexity alone results in about 20% of the explosions overlapping 40% of the earthquakes, compared with the third moment of frequency discriminant for which 20% of the explosions overlap about 15% of the earthquakes. Taking the data as a two dimensional discriminant (Fig. 1), however, a complete separation of the explosion population from the earthquake population is obtained when referenced to the arbitrary decision line $\text{TMF} = 13 + 0.25 \times C$. For one

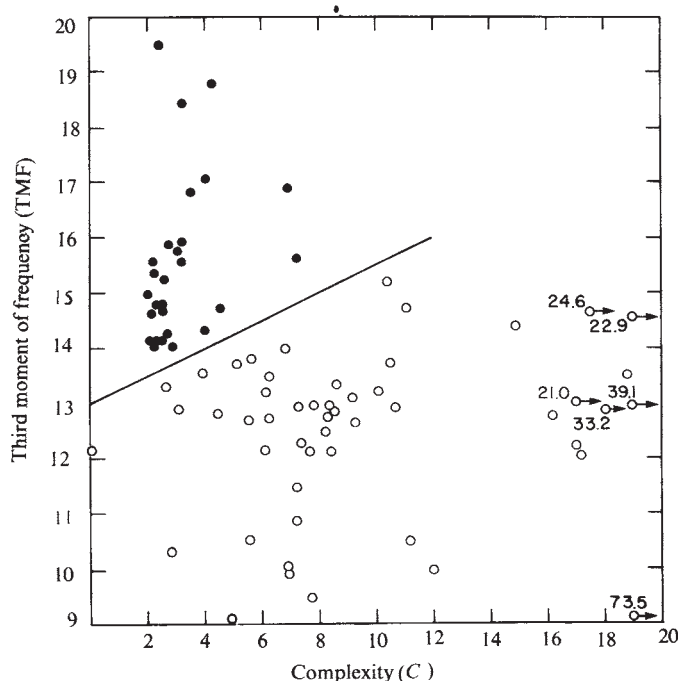


Fig. 1 Average third moment of frequency against phased beam complexity for a population of twenty-seven Eurasian explosions (●) and fifty-one shallow Eurasian earthquakes (○).

earthquake, C was set to zero as the data file was too short for the calculation of C . The orthogonal distances D of the data points from this line as a function of magnitude m_b (Fig. 2) shows D to be independent of m_b with the possible exception that the smaller explosions appear to be further from the decision line than the larger explosions. The value of D in this population has not been found to be a function of the seismic region or test area and for the earthquakes ($h < 50$ km) D has not been found to be a function of depth.

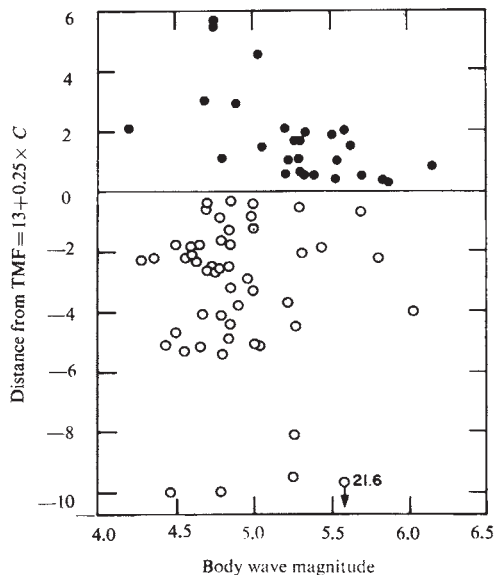


Fig. 2 Distance from the arbitrary decision line $TMF = 13 + 0.25 \times C$ as a function of body wave magnitude for a population of twenty-seven Eurasian explosions (●) and fifty-one shallow Eurasian earthquakes (○).

The empirical conclusion is that the earthquakes, which contain the higher frequencies, are sufficiently more complex than the corresponding explosions, so that complete separation is maintained using the array data from the regions of the five test sites. Similarly the more complex explosions such as those at Novaya Zembyla maintain a sufficiently high frequency character in the first 6.4 s.

As this study has not shown any need to regionalize the data within Asia, data are being collected from other test sites and seismic areas within the third seismic zone from Yellowknife. Extensive testing will be necessary to show whether this short period discriminant will continue to hold on a worldwide basis, or for larger event populations. If it does, it will provide a positive discriminant for explosions with yields smaller than 10 kT in hard rock.

I thank P. W. Basham, K. Whitham and E. Manchee for discussions.

F. M. ANGLIN

Seismology Division,
Earth Physics Branch,
Department Energy, Mines and Resources,
Ottawa K1A 0E4

Received July 16, 1971

- ¹ Evernden, J. F., *J. Geophys. Res.*, **74**, 15, 3828 (1969).
- ² SIPRI, *Seismic Methods of Monitoring Underground Explosions* (Almqvist and Wiksell, Stockholm, 1969).
- ³ Lincoln Laboratory, Massachusetts Institute of Technology; Technical Note, 1968-8, 1969-24; Semiannual Technical Summary (June 30, 1968), (December 31, 1968).
- ⁴ Weichert, D. H., Manchee, E. B., Whitham, K., *Geophys. J. Roy. Astron. Soc.* **13**, 277 (1967).
- ⁵ Lincoln Laboratory, Massachusetts Institute of Technology, Semiannual Technical Summary (December 31, 1969).
- ⁶ Weichert, D. H., *Z. Geophysik* (in the press).

Black Holes as Agents of Magnetic Fields

THE recent discovery of pulsating X-ray sources¹ has led to considerable speculation that these objects might be black holes. So far no mechanisms have been seriously proposed to explain how pulsed X-rays could be produced by black holes, but it is likely that any such mechanism would require the presence of a magnetic field. I wish to point out that it is highly unlikely that the magnetic field of a black hole could ever be large enough to produce observable effects.

It is well established that sufficiently massive stars must undergo complete gravitational collapse after they have exhausted their thermonuclear energy. All available theoretical evidence² suggests that the final state of this collapse is a Kerr-Newman black hole. In particular, there is considerable evidence (refs. 2-5 and unpublished results of R. Price and myself) that the only electromagnetic field a black hole can have is the Kerr-Newman electromagnetic field. The Kerr-Newman electromagnetic field differs substantially from an ordinary magnetic dipole field, but its magnetic dipole moment, given by the formula⁶,

$$\mu_K = |q|J/Mc$$

is a convenient indicator of the strength of the field. Here q is the charge of the black hole, J is its angular momentum, M is its mass, and c is the speed of light. One can obtain a firm upper limit on the value of μ_K from the following simple considerations: for a proton, $q_p/(\sqrt{G} M_p) \approx 10^{18}$ and for an electron $|q_e|/(\sqrt{G} M_e) \approx 10^{21}$, where G is the gravitational constant. If an uncollapsed star had a net positive charge $q > 10^{-18} \sqrt{G} M$ or a net negative charge $|q| > 10^{-21} \sqrt{G} M$, the electrostatic repulsive force on the excess protons or electrons would be greater than the gravitational attraction, and thus the excess charges would be driven away. Thus an uncollapsed star must have $|q|/(\sqrt{G} M) < 10^{-18}$. (On the other hand, a net positive charge somewhat less than this upper limit can be expected to occur since the less massive and faster moving electrons escape from the star more readily than protons.) No mechanism is known by which significant charge separation could occur during gravitational collapse; the above argument about the driving away of excess charge applies here also. Thus, the maximum value of $|q|/(\sqrt{G} M)$ that a Kerr-Newman black hole resulting from gravitational collapse could reasonably be expected to have is 10^{-18} . In fact, even if a black hole could be produced with $|q|/(\sqrt{G} M)$ somewhat greater than 10^{-18} , selective accretion of charged particles would be likely to lower the value of $|q|/(\sqrt{G} M)$ to less than 10^{-18} in a short time. In addition, for a Kerr-Newman black hole, one must have

$$\left(\frac{GM}{c^2}\right)^2 \geq \left(\frac{\sqrt{G}q}{c^2}\right)^2 + \left(\frac{J}{Mc}\right)^2$$

(This relation is necessary for the existence of an event horizon.) Thus one obtains

$$\mu_K < 10^{-18} G^{3/2} M^2 / c^2$$

or, in a more readily recognizable form,

$$\mu_K < 10^{-8} (M/M_\odot)^2 \mu_E$$

where μ_E is the magnetic dipole moment of the Earth. Even this small an electromagnetic field could still have an effect on the motion of charged particles near the event horizon. But it seems highly unlikely that this magnetic field could be large enough to produce any astrophysically observable effects.

ROBERT M. WALD

Joseph Henry Laboratories,
Princeton University,
Princeton, New Jersey 08540

Received June 22, 1971.

¹ Oda, M., Gorenstein, P., Gursky, H., Kellogg, E., Schreier, E., Tanabbaum, H., and Giacconi, R., *Astrophys. J. Lett.*, **166**, L1 (1971).

² Wald, R. M., *Phys. Rev. Lett.* (in the press).

³ Anderson, J. L., and Cohen, J. M., *Astrophys. Space Sci.*, **9**, 146 (1970).

⁴ Cohen, J. M., and Wald, R. M., *J. Math. Phys.* (in the press),
⁵ de la Cruz, V., Chase, J. E., and Israel, W., *Phys. Rev. Lett.*, **24**, 423 (1970).

⁶ Carter, B., *Phys. Rev.*, **174**, 1559 (1968).

Map of the Galactic Nucleus at 10 μm

We have scanned the centre of our galaxy at a wavelength of 10 μm with a beam 5.5 arcsec in diameter. The scans have been combined to produce the map shown in Fig. 1. At 10 kpc, commonly assumed to be the distance to this region, 5.5 arcsec corresponds to 0.3 pc.

These observations were taken during the day on February 15, 1971, with the NASA 61 inch telescope at the Catalina Observatory. Because of the imprecision in the drive of the telescope, the scans could not be located accurately in declination; the telescope also drifted slowly in right ascension. These problems were overcome by comparing the scans with some made earlier by one of us (F. J. L.) and D. E. Kleinmann with the Steward Observatory 90 inch telescope. Minor distortions may persist in the map. All the features described below stand out clearly above the noise, which has a peak-to-peak value of about $6 \times 10^{-17} \text{ W/m}^2 \text{ Hz ster.}$

Four sources can be distinguished and are numbered in Fig. 1. An additional broad zone of emission can be seen extending to the north. Sources 2 and 4 are partially resolved and have diameters of about 0.5 pc. There are indications on the individual scans that source 2 would break up into two or three sources at higher resolution. Source 1 is unresolved and must be less than 0.2 pc in diameter. Source 3 may also be unresolved with our beam, although the ratio of signal to noise is inadequate to permit a strong statement.

We have also made observations at 2.2, 5 and 22 μm . A detailed description of all this work will be published after we have analysed it more fully. We have located the "point-like" source discovered at 2.2 μm by Becklin and Neugebauer¹ relative to our map; to within the accuracy of the measurements, its position agrees with that of Source 3. At 5 μm , scans across the complex of sources have been made at five declinations. Although these scans have a relatively low ratio of signal to noise and are insufficient in number to construct a map, they do show the same general features as the 10 μm map.

At 22 μm , we have constructed a map with the same resolution and nearly the same ratio of signal to noise as the 10 μm map. The noise for this map has a peak to peak value of about $1.8 \times 10^{-16} \text{ W/m}^2 \text{ Hz ster.}$; at this level, sources 2 and 3 were not detected. Source 1 stands out prominently. It lies near the centre of an extended source which has a radius of about 20 arcsec.

Two field stars are plotted in Fig. 1 and labelled A and B. Additional scans at 10 μm enabled us to locate our map relative to field star A. The coordinates of this star were measured from a Palomar Sky Atlas plate, allowing us to determine the absolute position of the 10 μm sources to within about 2 arcsec. (A finding chart for the galactic centre has been published by Spinrad *et al.*² (their Fig. 1.) Star A is 2.5 mm south and 2.0 mm east of the indicated position of the galactic centre on this chart and star B is 1.5 mm north and 1.0 mm west. It should be noted that the position of the 10 μm sources differs somewhat from the one indicated on the published finding chart.)

From our map we estimate the total flux from the galactic centre to be $480 \pm 50 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$ at 10 μm . The 22 μm

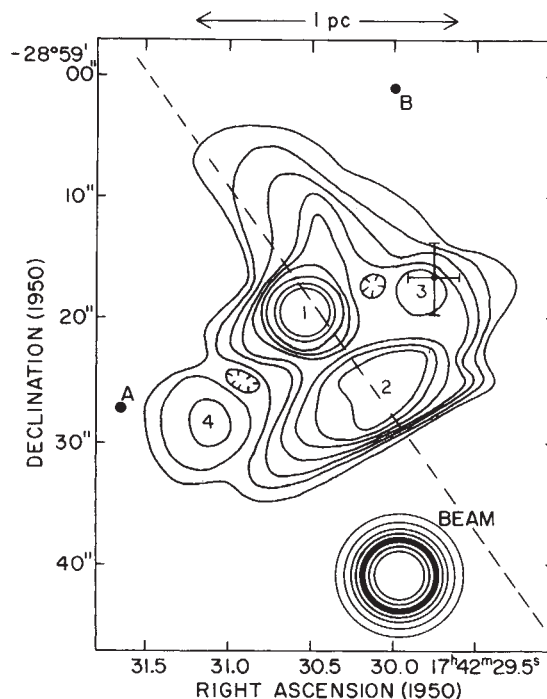


Fig. 1 Map of the galactic centre at 10 μm . The peak-to-peak noise of the observations is $6 \times 10^{-17} \text{ W/m}^2 \text{ Hz ster.}$ The contours are at levels of 1, 2, 4, 6, 8, 10, 12, 16 and 20 times this value. The brightest 10 μm sources are numbered in order of decreasing maximum flux. The beam used for these observations has a full width at half maximum of 5.5 arcsec. It is shown with the same contour levels as are used for the map. The location of the 2.2 μm "point-like" source is indicated by a set of error bars which correspond to the uncertainty in our measurement of its position. The two most prominent field stars have been plotted and labelled A and B. The direction of the galactic plane is indicated by the dashed line.

map indicates that the flux at this wavelength is $2,300 \pm 250 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$, of which about 60% or $1,400 \pm 150 \times 10^{-26} \text{ Hz}$ would fall within a field of view 25 arcsec in diameter. These results are in reasonably good agreement with earlier measurements made with a 25 arcsec field of view of $550 \pm 60 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$ at 10 μm and $1,700 \pm 200 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$ at 22 μm (ref. 3).

Maps of the galactic centre at 2.2 μm have already been published¹. At this wavelength, the detected flux arises almost entirely from unresolved stars. Except for the "point-like" source, the brightest features at 2.2 μm would produce a flux of $0.2 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$ in a field of view the size of ours. At 10 μm , the flux from these stars should be an order of magnitude lower. Our peak to peak noise is about $3.5 \times 10^{-26} \text{ W/m}^2 \text{ Hz}$; thus, the general stellar background detected at 2.2 μm makes virtually no contribution to the 10 μm map. Becklin and Neugebauer⁴ have published the results of a scan across the galactic centre at 10 μm . Their scan seems to have passed near the centre of source 2.

The 10 μm sources fall within the region of maximum intensity of the extended 2.2 μm source, which is centred¹ at $\alpha(1950) = 17 \text{ h } 42 \text{ m } 30 \text{ s} \pm 1 \text{ s}$ and $\delta(1950) = -28^\circ 59' 40'' \pm 0'.1$. Their position also agrees with that of the radio source Sgr A⁵, $\alpha(1950) = 17 \text{ h } 42 \text{ m } 30.6 \text{ s} \pm 1 \text{ s}$ and $\delta(1950) = -28^\circ 59' 14'' \pm 15''$.

In spite of the consistency of these positions, significant ambiguities remain if one looks closely at the data. Our 22 μm map suggests that source 1 lies at the galactic centre. On the other hand, a scan across the maximum of the 2.2 μm extended source, published by Becklin and Neugebauer⁴, places the greatest density of stars near source 2. The 2.2 μm "point-like" object can be identified with source 3; if it is, the 2.2 μm scan must be shifted in a direction which improves the correspondence with source 2. The existence of source 2

indicates that objects considerably cooler than normal stars lie very near the maximum 2.2 μ m intensity. The contribution of these objects to the 2.2 μ m flux is not known, leaving the position of the point of maximum stellar density uncertain.

Resolution of 6 arcsec at radio frequencies (D. Downes, personal communication) reveals that Sgr A consists of two objects, one to the east and one to the west of our position for the infrared sources. The published position reflects the contribution of both sources; its agreement with the infrared positions seems to be coincidental.

The first measurements of the discrete sources in the galactic nucleus were made by Low and D. E. Kleinmann⁶. These data were essential to our present efforts. We thank A. A. Hoag for giving us a plate from the Palomar Atlas and E. Roemer for assistance in carrying out the measurements.

G. H. RIEKE
F. J. LOW

University of Arizona,
Tucson, Arizona

Received July 29, 1971.

¹ Becklin, E. E., and Neugebauer, G., *Astrophys. J.*, **151**, 145 (1968).

² Spinrad, H., Liebert, J., Smith, H. E., Schweizer, F., and Kuhl, L. V., *Astrophys. J.*, **165**, 17 (1971).

³ Low, F. J., Kleinmann, D. E., Forbes, F. F., and Aumann, H. H., *Astrophys. J. Lett.*, **157**, L97 (1969).

⁴ Becklin, E. E., and Neugebauer, G., *Astrophys. J. Lett.*, **157**, L3 (1969).

⁵ Maxwell, A., and Taylor, J. H., *Astrophys. Lett.*, **2**, 191 (1968).

⁶ Low, F. J., Semaine d'Etude on the Nuclei of Galaxies, Vatican City, April, 1970.

BIOLOGICAL SCIENCES

Adaptive Significance of a Biased Sex Ratio in *Orchestia*

IN ecological studies of the marine supralittoral genus *Orchestia* (Crustacea, Amphipoda) many populations have been found that have biased sex ratios (ref. 1 and unpublished results of D. J. W.). Thelygeny has also been found in some of the terrestrial species of *Orchestia*², and in the sublittoral, species, *Orchestia gammarella* (Pallas), both amphogenous and thelygenous populations have been found³. I have investigated the possibility that biased sex ratio populations have an adaptive advantage in altering reproductive potential and the phylogenetic mechanism controlling its origin.

The annual reproductive potential (R) of an amphipod population may be expressed as:

$$R = X(bnp)$$

where X is the absolute number of adults, b the mean brood number, n the estimated number of broods in a season and p the percentage of females in the adult population. A theoretical consideration of the effect of sex ratio on R , ignoring physical and biological environmental factors and assuming b and n to be constants, is shown in Fig. 1. Unequal distribution of the sexes within a population would increase or decrease R by nearly 100%, provided each female could mate successfully. There are three categories of reproductive potential. First, there are the predominantly arrhenogenous populations in which R depends on the actual proportion of females and the intensity of male competition for mates. Second, there are those with a high or low proportion of females where mating competition is absent and the actual proportion of females controls R . And third, there are predominantly thelygenous populations where R depends on the actual proportion of

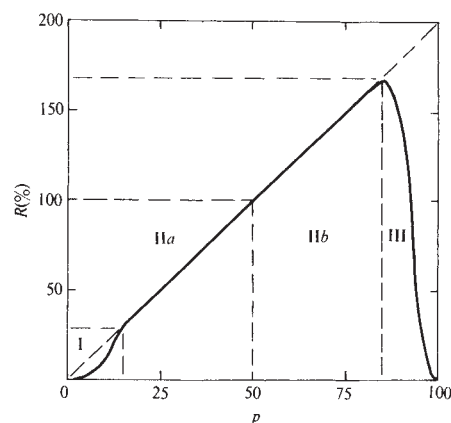


Fig. 1 Theoretical relationship between sex ratio and reproductive potential in *Orchestia*.

males available to mate with females and possibly on the intensity of female competition for mates.

In *Orchestia*, a biased sex ratio may occur when the population exhibits genetically controlled monogeny, when cross fertilization is the rule, parthenogenesis is absent, males are polygamous and when only one male can inseminate a female. Although these conditions occur in many of the known species of *Orchestia*, the exact genetic control mechanisms of sex determination and monogeny remain unknown. It has been suggested that in *O. gammarella* a polygenic system is involved⁴. There are three selective mechanisms, each operated by a different physical or biological environmental condition, which may determine the degree of sex bias in a given population. These are selective mortality, unequal fertility potential of monogenous parents and selective mating. Laboratory experiments have demonstrated selective female mortality and that the reproductive potential of *O. mediterranea* A. Costa, morph *mediterranea* is reduced when substrate and food salinity is low⁵. *O. mediterranea* is a polymorphic species that, near the limit of its estuarine habitat, is replaced by morph *aestuariensis* which differs from morph *mediterranea* in dorsal pigment patterns, the incidence of intersexes and in having a high proportion of reproductive females¹. The reproductive strategy of morph *aestuariensis* may be to maximize R (Type IIb) in the stringent estuarine conditions. It has also been shown¹ that crosses between *mediterranea* and *aestuariensis* have low fertility potential, although the reciprocal cross is fertile. In *O. gammarella*, discrete amphogenous, thelygenous or arrhenogenous colonies, apparently unrelated to physical environmental conditions, are known (ref. 3 and unpublished results of D. J. W.), although irregular fluctuations in the latter do affect optimum density conditions. In *O. gammarella* population density stresses could select biased sex ratio producing parents in isolated colonies. The mechanism would supplement the effects of density dependent regulation and provide a population homeorheostat that would maximize (Type IIb) or minimize (Type IIa) R (Fig. 2) according to the optimum permitted by the environment. An optimum R value would enable a wider range of niches to be occupied without stressing individual females. But whether natural or sexual selective processes are involved in this process is not known.

In relatively static populations which are part of a large inter-breeding group and integrated by sophisticated behavioural homeostatic mechanisms, as with many vertebrates, the stricture of Haldane⁶, that natural selection cannot influence population density, applies and the increased or decreased production is wasted. The situation in *Orchestia* populations was thought to deviate from this theory, the phenomenon of sex ratio bias being limited to demes where the population is integrated by primitive social behaviour and population regula-

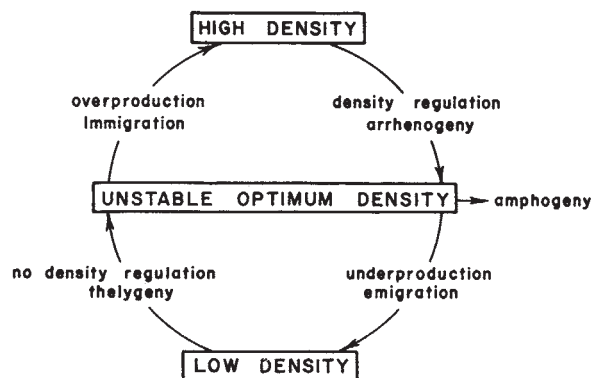


Fig. 2 Theoretical population homeorheostat in *Orchestia gammarella*.

tion methods, and/or where populations occur in discreet inbreeding colonies and are subject to irregular environmental fluctuations affecting optimum density. Sex bias would also cause a loss of genetic diversity⁷ and it was expected that where environmental conditions were stable there would be a tendency for the population to maintain an unbiased sex ratio.

The gene controlling fitness to survive or breed may either be linked to the one that controls sex expression or it may be the same one. If a polymorphic supergene of this type does control sex then the possibility of group selection as an explanation of biased sex ratio is unnecessary in *Orchestia*. It has been supposed that natural selective or sexual selective forces operating at the individual level are involved in sex ratio bias and thus monogeny need have no selective value for the individual, although as has been demonstrated, it may have considerable adaptive value in regulating reproductive strategy at the population level.

I thank A. M. Gower, N. J. Poole, V. Zitko and colleagues at St Andrews for helpful discussions.

D. J. WILDISH

Fisheries Research Board of Canada,
Biological Station, St Andrews,
New Brunswick

Received March 11, 1971.

¹ Wildish, D. J., *Crustaceana*, **19**, 113 (1970).

² Hurley, D. E., *Pacif. Sci.*, **13**, 107 (1959).

³ Charniaux-Cotton, H., *Ann. Sci. Nat. Zool. Biol.*, **19**, 411 (1957).

⁴ Ginsberger-Vogel, T., *Ann. Biol.*, **9**, 201 (1970).

⁵ Wildish, D. J., *J. Exp. Mar. Biol. Ecol.*, **5**, 276 (1970).

⁶ Haldane, J. B. S., *New Biol.*, **15**, 9 (1953).

⁷ Kalmus, H., and Smith, C. A. B., *Nature*, **186**, 1004 (1960).

Stereoscopic Vision in the Cat

STEREOPSIS is the perception of relative depth by means of the disparate retinal stimulation which occurs when horizontally separated eyes are fixed on a common point. The stimulus conditions which produce stereopsis in humans have been studied thoroughly, but only recently has it been possible to study the neural mechanisms¹⁻⁴. We have been able to demonstrate stereoscopic vision in the cat, a result which complements neurophysiological data demonstrating the presence of cells in cat visual cortex with receptive field geometry which is specialized for the detection of retinal disparity^{1,2}.

To test for stereopsis we used conditioned suppression⁵⁻¹², whereby an animal trained to emit a frequent response will inhibit or suppress that response on perception of a stimulus which signals a forthcoming noxious event. First, the cat was

trained to lick a water-filled tube continuously for a small food reward. During these training sessions the cat viewed three real vertical rods, arrayed in the same frontal plane, each subtending 18° and separated by 9° 33'. The rods, enclosed in a viewing tunnel, appeared as black lines seen against a back-illuminated homogeneous white projection screen, whose luminance was 0.32 foot Lamberts. Next, a conditioned suppression response was established, based on a discrimination between two real-depth positions of the centre rod. Near depth was introduced by moving the centre rod toward the cat, resulting in an angular disparity of 1° 2' relative to the stationary side rods; 5 s later a 1 s shock was delivered through the grid floor of the restraint box and then the rod returned to its original position. Far depth was introduced by moving the centre rod away from the cat, resulting in an angular disparity of 48', and returned to its original position 5 s later; no shock occurred during this kind of trial. Five near depth and five far depth counter-balanced trials were presented daily. Formation of the differential discrimination was clearly shown by the suppression of licking in the presence of near depth and the continuation of licking in the presence of far depth.

Once the discrimination between the near and far depth positions of the centre rod had been established, the cat was required to make the same discrimination for a rod-like shadow that appeared to move to near and far positions in stereoscopic depth. To produce this effect, a stereoscopic shadow caster was used, in which two point sources slightly separated horizontally cast the shadows of three vertical rods on a rear-projection screen. The shadows stimulated the eyes separately by means of chromatic filters (Wratten No. 58 green and No. 25A red) at the point sources and matching filters laminated to contact lenses placed in the cat's eyes. The geometric principles of shadow casting and the use of chromatic filters for dichoptic stimulation are established techniques, the theory and practical applications of which are well known^{13,14}. When the point source filters and eye filters are matched left and right and the centre rod is closer to the point sources, the resulting virtual image of the centre rod appears closer to the observer than the images of the side rods (crossed disparity). When the point source filters are interchanged, the virtual images are transposed and the centre rod appears farther from the observer (uncrossed disparity).

Stereoscopic depth discrimination was tested for 7 successive days, with four near depth trials (accompanied by shock) and four far depth trials each day. The centre rod was moved a distance which produced an angular disparity of 1° 9' (adjusted for the cat's interpupillary distance) between the centre shadow and the side shadows. Numerous blank trials, consisting of an interchange of the point source filters without movement of the centre rod, were included in the schedule; these trials produced no changes in the cat's behaviour. On all trials the eyes of the cat were observed closely, but neither eye closure nor head movement was observed.

In Fig. 1, performance on the last 3 days of real depth discrimination and on the 7 days of stereoscopic depth discrimination is expressed in terms of a suppression ratio ($W/S + W$, where W refers to the number of licks during the 5 s between rod movement and shock, and S refers to the number of licks during the 5 s preceding rod movement). There is little overlap between the distributions of near and far stereoscopic depth suppression ratios during the 7 days, a significant difference ($P < 0.01$) as measured by the non-parametric matched-pairs signed-ranks test.

It is not likely that the cat based its discrimination on a cue other than stereopsis. The intrinsic nature of the shadow caster precludes most cues. Control trials eliminated the operation of a potential compound cue based on filter interchange and lateral movement of the shadows. It is logically possible that the discrimination could be made if the cat could discern which shadow stimulated each eye (that is, possessed

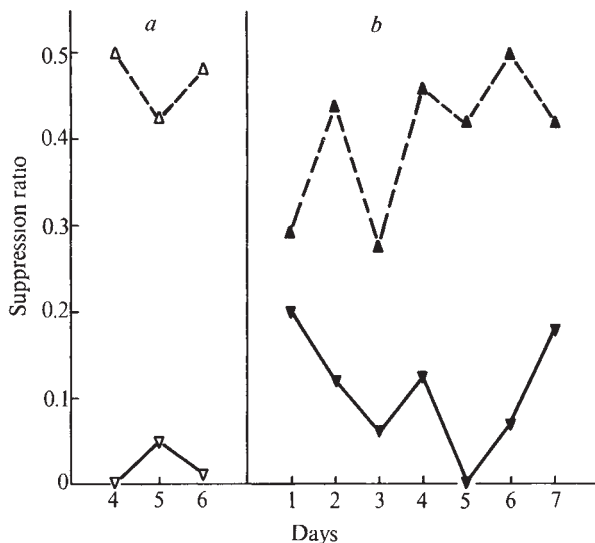


Fig. 1 Mean suppression ratios for (a) the last 3 days of real depth training and (b) 7 days of stereoscopic depth presentations. Δ , Backward; ∇ , forward; $\blacktriangle$, uncrossed disparity; $\blacktriangledown$, crossed disparity. Crossed disparity equals near stereoscopic depth; uncrossed disparity equals far stereoscopic depth. A ratio of 0.5 indicates no suppression; as ratios approach 0 suppression increases.

eye-of-origin information). Yet considerable research on humans indicates that eye-of-origin discriminations are virtually impossible¹⁵; humans who viewed our display could not make such discriminations. We think it unlikely that cats could readily do it when humans cannot.

Taken together, all considerations suggest that the cat possesses stereopsis and used it in making the conditioned suppression discrimination, an interpretation consistent with the neurophysiological data.

In this experiment a moving stimulus of large disparity was deliberately chosen to attract the cat's attention and, in general, to maximize the likelihood of demonstrating stereopsis. Now that the evidence suggests that the cat has some capacity for stereopsis, it should be possible to measure both the minimum and maximum limits of stereoscopic vision. Further, it may be possible to determine the ease with which cats can utilize the disparity information in random element stereograms^{16,17}.

We thank Drs R. Stroud, R. Althiasar and J. Greene for fitting, designing and constructing contact lenses. This study was supported in part by a US Public Health Service grant and a National Institute of Mental Health predoctoral fellowship.

ROBERT FOX
R. RANDOLPH BLAKE

Department of Psychology,
Vanderbilt University,
Nashville,
Tennessee 37203

- ¹³ Gregory, R. L., *Nature*, **203**, 1407 (1964).
- ¹⁴ Lee, D. N., *Vision Res.*, **9**, 145 (1969).
- ¹⁵ Templeton, W. B., and Green, F. A., *Quart. J. Exp. Psychol.*, **20**, 200 (1968).
- ¹⁶ Julesz, B., *Science*, **145**, 356 (1964).
- ¹⁷ Anonymous neurophysiological correspondent, *Nature*, **216**, 1067 (1967).

Relation between Prostaglandin E₂ and Adrenaline Reversal in the Rat Uterus

THE inhibitory response of the rat uterus to adrenaline in the presence of isoprenaline inhibition can be reversed in induced oestrus but not in the progestational state, and it has been postulated that a motor substance accumulates during the isoprenaline inhibition and is released on addition of adrenaline¹. The same sequence of events has also been observed when adrenaline was used as the initial inhibitory agent. The motor response to adrenaline is even more readily obtained in the late pregnant uterus but not in dioestrus².

Attempts have now been made to detect and identify the postulated motor substance. Many motor substances can be found in extracts of tissues, and to avoid confusion with these it was decided to look for the substance in the bath fluid, on the assumption that it was being released. The late pregnant uterus was chosen as a possible source because it apparently releases a greater quantity of the substance than when it is in induced oestrus, and adrenaline was used as the initial inhibitory agent. The dioestrous uterus was used to detect the substance, because reversal of adrenaline does not occur in this state. Mature female Sprague-Dawley rats weighing about 200 g were used. The stage of cycle was assessed by vaginal smearing. Pregnancies were dated and the day of finding sperm was denoted day 1. The animals were used on days 19–22. Uterine horns or strips were set up in organ baths according to a method described before^{1,2}. Adrenaline (100 ng/ml.) was added to a normally contracting strip from a late pregnant uterus; immediate inhibition occurred, but after about 2 min reversal occurred and the preparation began to contract. The bath fluid was then removed and transferred to a bath containing a dioestrous horn which had been inhibited with adrenaline for the same time. Two minutes after addition of the fluid the preparation began to contract. The response was small at first, increased and then died away, after which the preparation remained inhibited (Fig. 1a). The bath fluid contained adrenaline, but this had no motor effect on the dioestrous uterus, and the result therefore shows that a motor substance was being released into the bath fluid from the pregnant strip.

The results of experiments with the appropriate blocking agents eliminated the possibility that the motor substance might be acetylcholine, histamine or 5-hydroxytryptamine. Among the remaining possibilities, it was decided to test for the presence of a prostaglandin. The bath fluid was removed and adjusted to pH 3 with N HCl and extracted twice with ethyl acetate. The extract was then evaporated to dryness on a rotary evaporator; the residue was dissolved in 1 ml. of distilled water and applied to the rat fundal strip³. When the active substance was present it was examined by thin-layer chromatography with the AI and AII solvent systems of Green and Samuelsson⁴, using unequilibrated tanks. In ten determinations all biological activity was associated with a substance with a mean R_F of 0.66 in the AI solvent system. This behaved in a manner identical to that of a reference sample of authentic prostaglandin E₂ (R_F value 0.66), and the AII system showed it to be E₂. For confirmation parallel bioassays of the separated substance were done on guinea-pig ileum and rabbit jejunum. Prostaglandin E₂ was found in amounts varying from 30 ng to 1.5 μ g in ten extracts. Prostaglandins

- ¹ Barlow, H. B., Blakemore, C., and Pettigrew, J. D., *J. Physiol.*, **193**, 327 (1967).
- ² Joshua, O. E., and Bishop, P. O., *Exp. Brain Res.*, **10**, 389 (1970).
- ³ Hubel, D. H., and Wiesel, T. N., *Nature*, **255**, 41 (1970).
- ⁴ Bough, E. W., *Nature*, **255**, 42 (1970).
- ⁵ Estes, W. K., and Skinner, B. F., *J. Exp. Psychol.*, **29**, 390 (1941).
- ⁶ Brady, J. V., and Conrad, D. G., *J. Comp. Physiol. Psychol.*, **54**, 149 (1961).
- ⁷ Kamin, L. J., *J. Comp. Physiol. Psychol.*, **53**, 128 (1960).
- ⁸ Hoffman, H. S., Selekmán, W. L., and Fleshler, M., *J. Exp. Anal. Behav.*, **9**, 659 (1966).
- ⁹ Sidman, M., Ray, B., Sidman, R., and Klinger, J., *Exp. Neurol.*, **16**, 377 (1966).
- ¹⁰ Hendricks, J., *J. Exp. Anal. Behav.*, **9**, 501 (1966).
- ¹¹ Shumake, S. A., Smith, J. C., and Tucker, D., *J. Comp. Physiol. Psychol.*, **67**, 64 (1969).
- ¹² Hefner, H., and Masterton, R. B., *J. Comp. Physiol. Psychol.*, **71**, 175 (1970).

E_1 , $F_1\alpha$ and $F_2\alpha$ were also looked for but were not found. No prostaglandin was found in bath fluids from ten control strips incubated in the absence of adrenaline. This eliminated the possibility that prostaglandin E_2 was liberated as a result of the manipulation to which the uterine tissue had been subjected.

The possibility remained that liberation of prostaglandin E_2 was purely incidental. Further experiments were therefore carried out with a reference sample of authentic prostaglandin E_2 . Addition in a concentration of 20 ng/ml. to the bath fluid containing a dioestrous uterus was followed by a response which closely resembled in its course that produced by the motor substance liberated into the bath fluid (Fig. 1*b*). From this it seems that the liberation of prostaglandin E_2 was directly connected with the reversal of the inhibitory action of adrenaline.

From these results we conclude that the principal adrenergic action of the catecholamines on the rat uterus is stimulation of β -receptors resulting in inhibition. This supports the assertions of Levy and Tozzi⁵. In the presence of catecholamines prostaglandin E_2 is liberated and exerts a motor effect previously ascribed to stimulation of α -receptors in the rat uterus. Previous work has revealed anomalies in the behaviour of these receptors, and it has been suggested that they should be referred to as E-receptors pending further clarification². It now seems likely that they are receptors for prostaglandin E_2 . Adrenaline reversal in the rat uterus can therefore be explained by an initial β effect which is an intrinsic property of adrenaline, followed by a liberation of prostaglandin E_2 which then gives the opposite effect which has hitherto been interpreted as reversal.

Liberation of prostaglandin by catecholamines has also been observed in other organs (spleen⁶, rat epididymal fat pad⁷ and diaphragm⁸) and reversal of adrenaline has been documented for one of these⁶. Table 1 shows that there is a correlation

between the reversal response and the effect of prostaglandin E_2 in organs other than the uterus. These responses are clearly identical, and there is therefore good reason for supposing that adrenaline reversal, wherever it occurs, is brought about by liberation of prostaglandin.

Table 1 Correlation Between the Adrenaline Reversal and Prostaglandin E_2 Responses in Various Preparations

Preparation	Adrenaline response	Adrenaline reversal response	Prostaglandin E_2 response
Rat uterus oestrogen dominated	Relaxation	Contraction	Contraction
Rat fundal strip	Relaxation	Contraction	Contraction
Gut	Relaxation	Contraction	Contraction ⁹
Splenic vascular smooth muscle	Contraction ⁶	Relaxation ⁶	Relaxation ⁶
Blood pressure	Rise	Fall ¹¹	Fall ¹⁰

The results also indicate that the liberation of prostaglandin E_2 in the rat uterus is hormone-dependent, and this aspect of the problem and the site of liberation will be the subject of a further communication. We wish to thank Dr John Pike, of the Upjohn Co., Kalamazoo, for prostaglandin E_2 .

ANNE TOTHILL

Department of Pharmacology,
The London Hospital Medical College,
London E1

LEON RATHBONE

Department of Zoology,
Birkbeck College

EVE WILLMAN

Department of Pharmacology,
School of Pharmacy,
London WC1

Received March 25; revised June 3, 1971.

- ¹ Tothill, A., *Nature*, **213**, 1230 (1967).
- ² Tothill, A., *Brit. J. Pharmacol.*, **29**, 291 (1967).
- ³ Vane, J. R., *Brit. J. Pharmacol.*, **12**, 344 (1957).
- ⁴ Green, K., and Samuelsson, B., *J. Lipid Res.*, **5**, 117 (1964).
- ⁵ Levy, B., and Tozzi, S., *J. Pharmac. Exp. Ther.*, **142**, 178 (1963).
- ⁶ Davies, B. H., Horton, E. W., and Withrington, P. G., *Brit. J. Pharmacol.*, **32**, 127 (1968).
- ⁷ Shaw, J. E., *Fed. Proc.*, **25**, 770 (1966).
- ⁸ Ramwell, P. W., Shaw, J. E., and Kucharski, J., *Science*, **149**, 1390 (1965).
- ⁹ Bergstrom, S., Eliasson, R., Von Euler, U. S., and Sjövall, J., *Acta. Physiol. Scand.*, **45**, 133 (1959).
- ¹⁰ Du Charne, D. W., and Weeks, J. R., *Nobel Symp. 2, Prostaglandins*, 173 (Almqvist and Wiksell, Stockholm 1966).
- ¹¹ Butterworth, K. R., *Brit. J. Pharmacol.*, **21**, 378 (1963).

Intrinsic Limitations in the Use of Electroantennograms to Bioassay Male Pheromones in Lepidoptera

BEHAVIOURAL assays for the male pheromone of *Trichoplusia ni* (Lepidoptera: Noctuidae) are difficult to carry out because the female apparently lacks an overt behavioural response and mating only occurs at night in very dim illumination. Grant used the electroantennogram (EAG) as an assay¹, but although this technique has developed into a useful tool to bioassay female sex pheromones in Lepidoptera, it does not seem to be a biologically meaningful approach to this receptor system.

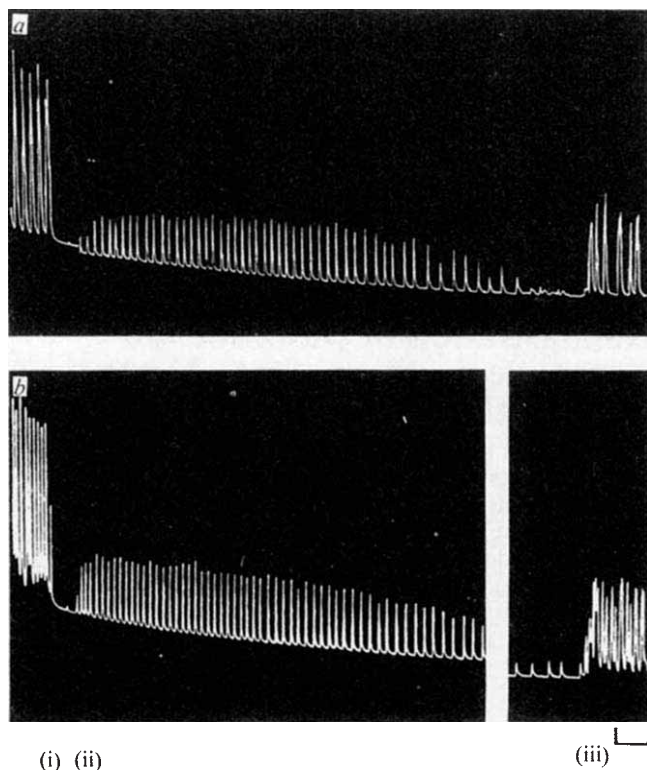


Fig. 1 Uterine horns from a rat in dioestrus. (i), Adrenaline was added to a final concentration of 0.1 μ g/ml. (ii), a, addition of bath fluid in which a strip of uterus from a pregnant horn had been treated with the same concentration of adrenaline; b, prostaglandin E_2 added to a final concentration of 0.02 μ g/ml. (iii) Wash. The break in b represents 52 min. Time marker, 5 min.

The EAG technique is based on the fact that the high sensitivity of many male Lepidoptera to female sex pheromone is achieved by a large proportion, up to 50%, of "odour specialist" receptors which respond at high sensitivity solely to the female sex pheromone². The summated response of these receptors over the length of the antenna gives the measured EAG d.c. potential change. Potential differences induced in neighbouring dendrites also summate over the length of the antenna to reinforce the effect. Only 200 pheromone molecules are needed to signal a behavioural response in males of *Bombyx mori*³ and a single molecule is sufficient to elicit firing in one of the specialist cells.

On the other hand, the male pheromones of Noctuidae are released in large quantity, about 0.1 mg in *Leucania pallens*⁴, very close to the female⁵ and so to the target receptors. With so much pheromone present at the receptors, there is little relevance, in relation to natural behaviour, to behavioural or EAG measurements of threshold responses. Grant¹ shows that the excized hairpencils of one male *T. ni* cause a depolarization of 1.0 mV in both male and female antennae, whereas tufts of hair from elsewhere on the male body do not evoke a response. The EAG, however, does not show that the male can distinguish the stimulus as a pheromone from the general olfactory spectrum of the species.

I have recorded EAGs in response to male scent in a number of Noctuidae. Standard procedures were followed using KCl filled glass pipette electrodes and immobilized whole moths rather than isolated heads or antennae. EAG responses to stimuli of conspecific male brushes, to cross-specific brushes and to stimulus equivalent extracts did not differ significantly for each species regardless of whether recording from male or female antennae. Levels of EAG depolarization varied between species, but EAGs within a species were consistent in response to the stimulus of brushes from all other species used. The values obtained were: *Leucania pallens* (L.), $-3.4 (\pm 0.3)$ mV; *Leucania conigera* (Schiffermueller), $-3.3 (\pm 0.3)$ mV; *Apamea monoglypha* (Hufnagel), $-2.0 (\pm 0.2)$ mV; *Phlogophora meticulosa* (L.), $-2.6 (\pm 0.2)$ mV; and *Mamestra persicariae* (L.), $-2.1 (\pm 0.3)$ mV. Stimulus equivalent quantities of compounds isolated from the male brushes and various related standard laboratory compounds elicited EAGs indistinguishable from those produced by the male brushes for a range of dilutions. The isolated compounds are low molecular weight aromatics, monoterpenes and simple carboxylic acids⁶. Positive responses to such a range of compounds are not considered to be artefacts since many other related compounds do not evoke EAG responses.

Grant shows that the female sex pheromone of *T. ni* gives an EAG in the female antenna of about the same size (1 mV) as that produced in both sexes by the male brushes, whereas the female sex pheromone evokes a larger (3 mV) response in the male antenna. Grant uses Callahan's evidence of sexual dimorphism in the antennae of *Heliothis zea*⁷ to suggest that perception by morphologically distinct receptors on male and female antennae explains the difference. Jefferson *et al.*⁸, however, have shown that there are negligible differences in the number and distribution of sensillae between the sexes in *T. ni*. There is greater variation between individuals and between antennae on the same moth. This does not support Grant's hypothesis and it seems more likely that one type of sensilla has become functionally differentiated and very narrowly specialized in the female alone.

The male releases pheromones so close to the female that there has been no selective pressure to develop a receptor which responds preferentially to minute levels of one substance. The receptors which do respond are the "odour generalists"², each of which has a different odour spectrum signalled, for example, by latency of rate of firing, or by speed of adaptation to a given odour. Signals from the odour generalists are deciphered by the central nervous system, and it would require an analogous

detailed analysis of nerve spikes, or single cell recordings, to separate clearly the response to one odour or component thereof from that to another. EAG responses can be obtained from these receptors but, as a similar number of receptors respond to similar concentrations of many odorous compounds within the olfactory spectrum of the species, the summated EAG recordings are also similar.

EAGs show only that the male Noctuid pheromones can be perceived peripherally. For a more detailed knowledge of the way in which these chemicals stimulate or modify specific behaviour, there is no present alternative to using a behavioural assay. It is dangerous to use techniques based on response at low threshold levels by odour specialist cells to monitor processes based on neither. We still do not know what information is carried by nerve impulses, how it is integrated by the central nervous system, or even the difference between excitatory and inhibitory impulses. Where EAGs are used to investigate the behaviour of insects to olfactory stimuli, it is essential that they are closely correlated with at least laboratory behavioural studies and preferably field investigations.

I thank Drs D. Schneider and E. Priesner for very valuable discussions at Seewiesen and Drs R. T. Aplin, G. G. Grant, T. L. Payne and D. L. Wood for much helpful advice and criticism. The Anti-Locust Research Centre, London, allowed me initial use of a recording apparatus and Mr R. Yeadon gave help in its use.

MARTIN C. BIRCH

Department of Entomology,
University of California,
Berkeley, California 94720

Received April 25; revised June 8, 1971.

¹ Grant, G. G., *Nature*, **227**, 1345 (1970).

² Schneider, D., *Science*, **163**, 1031 (1969).

³ Kaissling, K.-E., and Priesner, E., *Naturwissenschaften*, **57**, 23 (1970).

⁴ Birch, M. C., *Trans. R. Entomol. Soc. Lond.*, **122**, 277 (1970).

⁵ Birch, M. C., *Anim. Behav.*, **18**, 310 (1970).

⁶ Aplin, R. T., and Birch, M. C., *Experientia*, **26**, 1193 (1970).

⁷ Callahan, P. S., *Res. Bull. Univ. Georgia, Coll. Agric. Exp. Stn.*, **54** (1969).

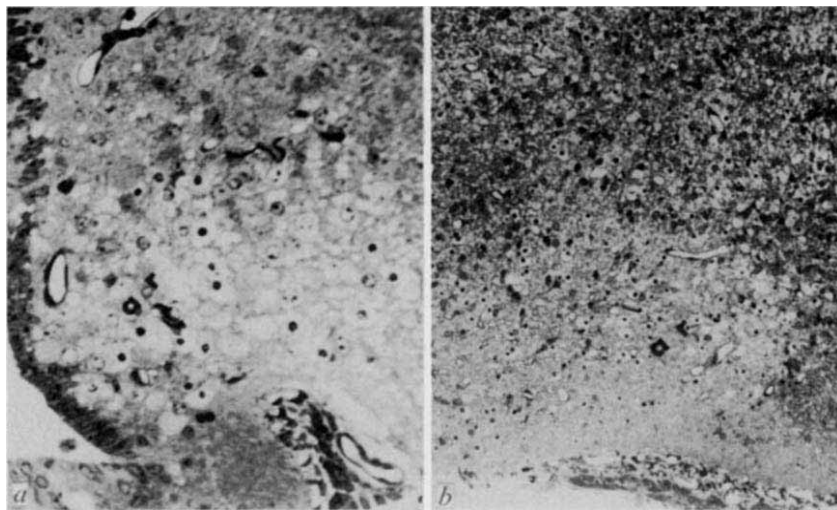
⁸ Jefferson, R. N., Rubin, R. E., McFarland, S. U., and Shorey, H. H., *Ann. Entomol. Soc. Amer.*, **63**, 1227 (1970).

Acute Effect of Oral and Subcutaneous Administration of Monosodium Glutamate on the Arcuate Nucleus of the Hypothalamus in Mice and Rats

RECENT reports by Olney that monosodium glutamate (MSG) can cause acute neuronal necrosis in the developing brain of mice¹⁻³, rats¹ and monkeys⁴, have been the subject of intensive debate⁵⁻¹¹. Some people^{9,11} deny that a single subcutaneous dose of monosodium glutamate can produce any brain damage in infant rats. Although Arees and Mayer¹⁰ confirm the existence of a necrotic lesion in both mouse and rat brain they restrict the pathology to "microglial" cells. None of these investigators attempted actually to replicate Olney's methods but used entirely different methods of preservation and staining.

Because any reports of this nature demand independent confirmation, a blind study was undertaken to determine both the presence or absence of and the consistency of an MSG effect on the developing hypothalamus of mice using the exact methods of administration of MSG, fixation and staining outlined by Olney² (Table 1). Further studies were completed in a similar experimental situation on Wistar rats using Olney's techniques as well as those of Adamo and Ratner⁹ (Table 1).

Fig. 1 *a*, Section through the hypothalamus of 10 day old Swiss albino mouse 5 h after oral administration of 1 mg/g of MSG, illustrating a distinct lesion in the region of the arcuate nuclei (AR). The neuronal elements are necrotic with a bloated cytoplasm and pyknotic nuclei ($\sim \times 650$). *b*, Section through the hypothalamus of 4 day old Wistar rat 5 h after oral administration of 4 mg/g of MSG, demonstrating a massive lesion involving the arcuate (AR) and ventromedial (VM) nuclei ($\sim \times 450$).



Monosodium glutamate was administered either orally or by subcutaneous injection. Sections for light microscopy were read independently by two of us who were unaware of the treatment schedule. Sections for electron microscopy were collected on slot grids and stained with uranyl acetate and lead citrate.

Table 1 Treatment Schedule and Fixation

Animal	Method of administration	Treatment		Perfusion method
		MSG	NACl/saline	
Mouse	Oral	9 (1 mg/g)	9	Olney
Rat	Oral	5 (4 mg/g)	4	Olney
Rat	Oral	4 (4 mg/g)	4	Adamo and Ratner
Rat	Subcutaneous	6 (4 mg/g)	4	Adamo and Ratner
Rat	Subcutaneous	2 (4 mg/g)	2	Olney
		2 (2 mg/g)		

Light microscopic examination of the brains of infant mice treated orally with monosodium glutamate at a dose of 1 mg/g revealed that all the animals developed lesions in the arcuate area of the hypothalamus and median eminence (Fig. 1*a*), with complete correlation between independent examiners. No evidence of cellular pathology was noted in the controls. Light microscopic examination of the brains of infant rats treated by either route of administration with either 2 mg/g or 4 mg/g of monosodium glutamate similarly demonstrated arcuate lesions in all treated animals (Fig. 1*b*) with an absence of cellular pathology in control animals. Lesions were much more extensive in rats than mice, probably reflecting either the higher dosage used or possibly the younger age of the rats.

The infant rat brains perfused with formalin were extremely fragile. Tissue preservation was not satisfactory for photography, but the affected areas could not be identified and were in the periventricular arcuate area. The lesion could be seen better when haematoxylin and eosin were used instead of cresyl violet.

Examination of both rat and mouse material with the electron microscope revealed necrosis of the neuronal components of the arcuate nucleus. Perikaryon swelling, loss of cytoplasmic density and organelles and nuclear pyknosis are evident. The neuronal nature of these cells, at least in the mouse, is demonstrated by the presence of synaptic contacts both on cell bodies and dendrites (Fig. 2). The presynaptic components, that is, vesicles and membrane density, remain intact, while there is dissolution of the postsynaptic structure in the involved cells.

A more recent report by Olney³ of studies using both light and electron microscopy, outlines the development of monosodium glutamate-induced lesions from acute neuronal necrosis through biological degradation and phagocytosis to an apparently normal residual cytology except for a markedly decreased population of neurones in the arcuate area. Thus, the photomicrographs of Adamo and Ratner⁹, who examined only adult brains by electron microscopy and found no evidence

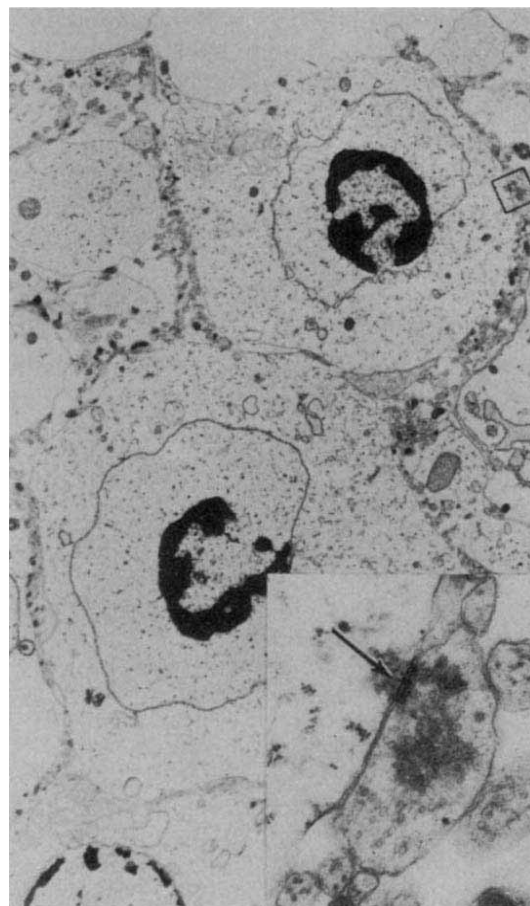


Fig. 2 Electron micrograph of degenerating neurones illustrating a dissolution of cytoplasmic components and alteration of the nuclear chromatin pattern. All the cytoplasmic structures have degenerated beyond recognition ($\times 12,000$). The inset on the lower right is a higher magnification of the area within the box. Apparently normal axon terminal with presynaptic vesicles and increased membrane density (arrow) in synaptic contact with the soma of a degenerating cell. A remnant of post-synaptic structure is evident ($\times 50,000$).

of cellular pathology at the ultrastructural level, are entirely consistent with Olney's¹.

The question remains why neither Adamo and Ratner⁹ nor Oser *et al.*¹¹ found any MSG effect in their light microscopic studies of infant brains of animals. One can only speculate that the method of fixation and staining used by Adamo and Ratner obscured the existence of the lesion. On the other hand, Oser *et al.*¹¹ used a minimal effective dose and did not examine the rats and mice until 24 h after insult. Olney³ has demonstrated that by 24 h after insult, all signs of oedema have disappeared and that necrotic cells are already being phagocytized. Oser *et al.*¹¹ reported finding occasional areas of small pyknotic neurones, large cytoplasmic vacuoles and scattered inflammatory cells. This is entirely compatible with a lesion late in evolution which would be expected if the insult was minimal.

The failure of Adamo and Ratner to find evidence of endocrine dysfunction is not surprising in face of their dose schedule. Olney¹ states that mice treated with a single dose of MSG are obese compared with control litter mates, but he does not allude to any reproductive dysfunction in this group. Reproductive dysfunction has been reported only in the group of animals treated with consecutive doses of MSG during infancy^{12,13}.

The interpretation by Arees and Mayer¹⁰ that the lesion produced by MSG is limited to "microglia" but totally spares hypothalamic neurones is puzzling. Most cells normally present in the arcuate nucleus are small neurones¹⁴. Olney³ has demonstrated, and we have confirmed, that most of the cells of the arcuate nucleus manifest extreme degenerative changes 5 h after MSG treatment even at the light microscopic level. The demonstration of synaptic involvement by Olney³ and in this article, as well as the previously documented nature of the effect of MSG in the retinae¹⁵⁻¹⁷, clearly points to a lesion involving neurones. These synapses were documented both on the soma and dendrites of damaged cells. Although there is considerable confusion surrounding the concept of "microglia", it is reasonably clear that axon terminals do not synapse on either the processes or the cell bodies of "microglia".

The purpose of this article is not to enter the debate on the safety of monosodium glutamate but merely to confirm or deny previous results in a scientifically valid manner. Using the exact experimental protocol described by Olney and Ho², and blind analysis of experimental and control specimens, we have confirmed their claim that 100% of infant mice treated orally with 1 mg/g of MSG develop hypothalamic lesions but that animals subjected to an equivalent oral dose of sodium chloride do not. The similar effect of MSG on the infant rat brain has been demonstrated by both the oral and subcutaneous route of administration. Finally, we believe that the exact duplication of experimental protocols, including age of animals, route of administration, doses and method of fixation, and careful interpretation of obtained results, in the light of carefully read previous reports, is necessary before valid comparisons can be made.

This work was supported in part by grants from the National Eye Institute. Animals were supplied by National Laboratory Animals, Missouri, and MSG was supplied by Sigma Chemical Company.

R. M. BURDE
B. SCHAIKNER
J. KAYES

Department of Ophthalmology
and Oscar Johnson Institute,
Washington University School of Medicine,
Saint Louis,
Missouri 63110

Received March 5; revised May 11, 1971.

¹ Olney, J. W., *Science*, **164**, 719 (1969).

² Olney, J. W., and Ho, O.-L., *Nature*, **227**, 609 (1970).

³ Olney, J. W., *J. Neuropathol. Exp. Neurol.*, **30**, 75 (1971).

⁴ Olney, J. W., and Sharpe, L. G., *Science*, **166**, 386 (1969).

⁵ Lowe, C. U., *Science*, **167**, 1016 (1970).

⁶ Zavon, M. R., *Science*, **167**, 1017 (1970).

⁷ Olney, J. W., and Sharpe, L. G., *Science*, **167**, 1017 (1970).

⁸ Subcommittee on Safety and Suitability of MSG and other Substances in Baby Foods (National Academy of Sciences—National Research Council, 1970).

⁹ Adamo, N. J., and Ratner, A., *Science*, **169**, 673 (1970).

¹⁰ Arees, E. A., and Mayer, J., *Science*, **170**, 549 (1970).

¹¹ Oser, B. L., Carson, S., Vogen, E. E., and Cox, G. E., *Nature*, **229**, 411 (1971).

¹² Knittle, J. J., and Ginsberg-Follner, F., *Amer. Pediat. Soc. Bull.* (program abstract, April 29, 1970).

¹³ Redding, R. W., and Schally, A. V., *Fed. Proc.*, **29** (1970).

¹⁴ Christ, J. F., in *The Hypothalamus* (edit. by Haymaker, W., Anderson, E., and Nanta, W.), 36 (Thomas, Springfield, Illinois, 1969).

¹⁵ Lucas, D. R., and Newhouse, J. P., *Amer. Med. Assoc. Arch. Ophthalmol.*, **58**, 193 (1957).

¹⁶ Cohen, A. I., *Amer. J. Anat.*, **120**, 319 (1967).

¹⁷ Olney, J. W., *J. Neuropathol. Exp. Neurol.*, **28**, 455 (1969).

Comments on "Neuromythology"

THE comments by Comfort on Dawkins's¹ memory theory struck a sympathetic chord, late though they were received here. We failed to detect any of the changes claimed² for the nucleus dentatus and found but a fraction³ of the neurones estimated. Counts covering newborn to presenescent age produced no age correlation in the human olivary nucleus⁴. The need for the provision of more cell counts stressed by Comfort was emphasized too by Blinkow⁵ at a recent meeting here. Neuronal population studies done by count are certainly more tedious than those based on general estimates. They render figures more like those of bankers than of economists. Delorenzo⁶ found no evidence for the loss of Purkinje cells in man reported by Ellis⁷. Thus not only for rodents, but in man too, the more recent studies based on larger material, fail to demonstrate loss of neurones with advancing age. Available information about the regular Purkinje cell layer is so contradictory, the multilayered cerebral cortex with cell number estimates ranging^{8,9} from 2.6×10^9 to 16.5×10^9 , that it is even less likely to provide unequivocal data easily. The billion range is not of variability but one of uncertainty. Side by side illustrations demonstrating age changes suffer from what may be called the Kara Mustapha fallacy*. Recent work by Dayan¹⁰ raises doubts about some more tenets of the neuropathology of ageing. Dawkins's model of memory helps to bring to light the narrow base supporting neurone loss as an uncontested axiom.

JOSEPH TOMASCH

Department of Anatomical Sciences,
Pahlavi University,
Shiraz,
Iran

Received May 17, 1971.

* Unsuccessful in his siege of Vienna, Kara Mustapha was presented with a silken rope by the sultan's emissary. Museum guides maintain that the skull on display in Budapest is of an earlier period than the one in the Vienna museum.

¹ Dawkins, R., *Nature*, **229**, 118 (1971).

² Hoepker, W., *Z. Alternsforsch.*, **5**, 256 (1951).

³ Heidary, H., and Tomasch, J., *Acta Anat.*, **74**, 290 (1969).

⁴ Moatamed, F., *J. Comp. Neurol.*, **128**, 109 (1966).

⁵ Blinkow, S. M., *Proc. Pahlavi Med. Congr.*, **2**, 181 (1971).

⁶ Delorenzo, E., *Z. Zellforsch. Mikr. Anat.*, **14**, 310 (1932).

⁷ Ellis, R. S., *J. Comp. Neurol.*, **32**, 1 (1920).

⁸ Pakkenberg, H., *J. Comp. Neurol.*, **128**, 17 (1966).

⁹ Haug, H., and Rebhahn, I., *Acta Anat.*, **28**, 551 (1956).

¹⁰ Dayan, A. D., *Acta Neuropathol.*, **16**, 85 (1970).

Sensitivity of Fish ATPases to Polychlorinated Biphenyls

In spite of widespread concern about the possible hazards of polychlorinated biphenyls (PCBs)¹⁻¹¹, little is known about the physiological action of these materials. We have therefore examined the effect of PCBs on the ATPase enzyme system, which is known to be sensitive to chemically similar compounds such as DDT¹²⁻¹⁸.

It is known that PCB ('Aroclor 1242') will cause oedema and depress growth when fed to chicks in large amounts (200 and 400 p.p.m. in diet), the rate of steroid metabolism is increased in liver homogenates from birds dosed with PCBs and that the materials stimulate oestrogenic activity in the rat uterus^{10,11}. We have examined the activity of the ATPase enzyme system in homogenized and fractionated tissues of the blue gill fish *Lepomis macrochirus*, because this system is sensitive to DDT and related compounds and because distinctive differences in effect have been found among cyclodiene-type chlorinated hydrocarbons.

Tissues were dissected, homogenized and fractionated according to the procedure reported by Koch¹⁸. The fraction used (B) contained mitochondria plus nerve endings. Each preparation was diluted appropriately, quick frozen in liquid nitrogen and stored at -20° C. ATPase activities were determined essentially according to the enzymatic procedure described before¹⁹⁻²¹. Proteins were determined by the method of Lowry *et al.*²². Reaction mixtures used are given in Table 1.

Table 1 Percentage Inhibition of Mg²⁺ ATPase from Muscle of Blue Gill Fish by PCB Preparations*

Concentration (p.p.m.)	'Aroclor 1221'	'Aroclor 1254'	'Aroclor 1268'	'Aroclor 5460'
0.03	18.7	20.6	37.2	31.9
0.33	29.9	30.6	47.1	42.9
3.33	65.4	73.1	60.3	—
10.00	80.4	84.3	67.5	—

A 3 ml. reaction mixture contained compounds at the following concentrations: 4.3 mM ATP, 5 mM Mg²⁺, 100 mM Na⁺, 20 mM K⁺, 135 mM imidazole buffer (pH 7.5), 0.19 mM NADH, 0.5 mM PEP, 0.02% BSA, approximately 9 U pyruvate kinase and 12 U lactic dehydrogenase, and 100 µl. homogenate fraction. Absorbance changes were measured at 340 nm for 15 min using a Beckman DU spectrophotometer with temperature controlled at 27° C in reaction mixture. When used, ouabain was present at 0.85 mM. Total ATPase activity was determined in the presence of Na⁺, K⁺ and Mg²⁺. Mg²⁺ ATPase activity was measured using same mixture plus ouabain. Na⁺-K⁺ ATPase activity is total activity minus Mg²⁺ ATPase activity^{20,21}. Ethanol was used as solvent for the 'Aroclor' except 'Aroclor 1268', where acetone was used. Between 1 and 5 µl. of the 'Aroclor' solutions was added to the enzyme reaction mixture by slowly releasing the solution from a Hamilton microsyringe into a rapidly stirred reaction mixture.

* Each value represents the mean from three separate preparations. Specific activities (µmol Pi mg⁻¹ protein h⁻¹; mean ± s.e.) for untreated muscle Mg²⁺ ATPase and Na⁺-K⁺ ATPase were 57.1 ± 2.2 and 2.3 ± 0.8 respectively. Low specific activity of Na⁺-K⁺ ATPase in muscle made accurate determinations of inhibition very difficult.

Three PCB preparations, 'Aroclor 1221', 'Aroclor 1254' and 'Aroclor 1268', and one polychlorinated terphenyl, 'Aroclor 5460' (Monsanto) were used. The last two digits of the name give the percentage chlorination of each compound, but the exact composition is not available. All compounds inhibited ATPases in low concentrations (less than 1 p.p.m.) with the most pronounced inhibition on Mg²⁺ ATPase of muscle (Tables 1 and 2). The higher sensitivity of muscle Mg²⁺ ATPase to PCB compounds is consistent with the high sensitivity of this enzyme to DDT¹³⁻¹⁵.

'Aroclor 1254' was the most effective inhibitor of ATPases, especially Mg²⁺ ATPase, from all tissues (Tables 1 and 2). 'Aroclor 1268', an effective inhibitor of muscle Mg²⁺ ATPase (Table 1), differed from the other PCB compounds in having minimal effects on Mg²⁺ ATPase activity from other tissues (Table 2). 'Aroclor 5460' seemed to be intermediate between 'Aroclor 1268' and the other 'Aroclor' compounds in its effect on Mg²⁺ ATPase activity (Table 2). The variation between tissues, most prominent with 'Aroclor 1268', indicates that the characteristics of Mg²⁺ ATPase differ depending on tissue source. A 30-40% inhibition of Na⁺-K⁺ ATPase measured only on brain and kidney preparations was consistently determined for all PCBs.

Table 2 Inhibition of Mg²⁺ ATPase and Na⁺-K⁺ ATPase from Tissues of Blue Gill Fish

Concentration (p.p.m.)	Percentage inhibition *							
	'Aroclor 1221' Mg ²⁺	Na ⁺ -K ⁺	'Aroclor 1254' Mg ²⁺	Na ⁺ -K ⁺	'Aroclor 1268' Mg ²⁺	Na ²⁺ -K ²⁺	'Aroclor 5460' Mg ²⁺	Na ²⁺ -K ²⁺
(A) Brain								
0.03	0.3	4.6	4.2	16.9	0.4	29.4	2.3	26.9
0.33	8.4	10.5	0.7	19.1	+1.5†	27.6	2.9	26.3
3.33	7.2	20.4	27.7	41.5	+14.0	43.6	2.5	38.9
10.00	36.0	46.4	57.9	45.5	8.3	39.4	11.2	42.7
(B) Kidney								
0.33	7.4	15.1	1.3	35.2	0.0	25.4	0.0	22.3
3.33	5.0	24.3	33.4	45.9	+7.9	40.1	4.5	25.2
10.00	42.8	43.3	60.1	51.7	+4.8	48.0	+1.4	33.8
(C) Liver								
0.03	2.2	—	3.9	—	—	—	—	—
0.33	2.6	—	19.4	—	+5.3	—	6.6	—
3.33	2.9	—	24.3	—	+5.1	—	9.3	—
10.00	35.8	—	53.0	—	+2.4	—	15.8	—
16.70	45.9	—	74.1	—	6.2	—	—	—

A, Specific activities for untreated brain, Mg²⁺ ATPase 25.2 ± 1.0, Na⁺-K⁺ ATPase 35.8 ± 2.1.

B, Specific activities for untreated kidney, Mg²⁺ ATPase 31.3 ± 1.3; Na⁺-K⁺ ATPase 42.0 ± 0.7.

C, Specific activities for untreated liver, Mg²⁺ ATPase 28.5 ± 3.0; Na⁺-K⁺ ATPase 2.8 ± 0.8; low specific activity of Na⁺-K⁺ ATPase in liver made accurate determinations of inhibition very difficult.

For reaction conditions see Table 1 except that temperature was 37° C.

* Each value represents the mean from three separate preparations.

† Plus values represent percentage activation of enzymatic activity.

In general, the PCB preparations do not reveal any obvious relationships between percentage chlorination and inhibitory effects. But the differences in action of 'Aroclor 1268' on Mg²⁺ ATPase from different tissues and the similarities in effects between PCBs and DDT require further study.

This research was supported by a grant from the US Department of Interior.

H. H. YAP
D. DESAIAH
L. K. CUTKOMP

Department of Entomology,
Fisheries and Wildlife,
University of Minnesota, St Paul

R. B. KOCH

Honeywell Research Center,
Hopkins, Minnesota 55343

Received January 11; revised May 24, 1971.

- ¹ Anderson, D. W., Hickey, J. J., Risebrough, R. W., Hughes, D. F., and Christensen, R. E., *Canad. Field Naturalist*, **83**, 89 (1969).
- ² Holmes, D. C., Simmons, J. H., and Tatton, J. O'G., *Nature*, **216**, 227 (1967).
- ³ Koeman, J. H., Ten Noever De Brauw, M. C., and De Vos, R. H., *Nature*, **221**, 1126 (1969).
- ⁴ Jensen, S., Johnels, A. G., Olsson, S., and Otterlind, G., *Nature*, **224**, 247 (1969).
- ⁵ Risebrough, R. W., Reiche, P., Peakall, D. B., Herman, S. G., and Kirven, M. N., *Nature*, **220**, 1098 (1968).
- ⁶ Duke, T. W., Lowe, J. I., and Wilson, jun., A. J., *Bull. Environ. Contam. Toxicol.*, **5**, 171 (1970).
- ⁷ Peakall, D. B., and Lincer, J. L., *Bioscience*, **20**, 958 (1970).
- ⁸ Biros, F. J., Walker, A. C., and Medbery, A., *Bull. Environ. Contam. Toxicol.*, **5**, 317 (1970).
- ⁹ Flick, D. F., O'Dell, R. G., and Childs, V. A., *Poultry Sci.*, **44**, 1460 (1965).
- ¹⁰ Lincer, J. L., and Peakall, D. B., *Nature*, **228**, 738 (1970).
- ¹¹ Bitman, J., and Cecil, H. C., *J. Agric. Food Chem.*, **18**, 1108 (1970).
- ¹² Koch, R. B., *J. Neurochem.*, **14**, 269 (1969).
- ¹³ Koch, R. B., *Chem.-Biol. Interact.*, **1**, 199 (1969/70).
- ¹⁴ Koch, R. B., Cutkomp, L. K., and Do, F. M., *Life Sci.*, **8**, 289 (1969).
- ¹⁵ Cutkomp, L. K., Yap, H. H., Cheng, E. Y., and Koch, R. B., *Chem.-Biol. Interactions* (1971).
- ¹⁶ Matsumura, F., Bratkowski, T. A., and Patil, K. C., *Bull. Environ. Contam. Toxicol.*, **4**, 262 (1969).
- ¹⁷ Matsumura, F., and Patil, K. C., *Science*, **166**, 121 (1969).
- ¹⁸ Koch, R. B., *J. Neurochem.*, **16**, 145 (1969).
- ¹⁹ Pullman, M. E., Penefsky, H. S., Datta, A., and Racker, E., *J. Biol. Chem.*, **235**, 3322 (1960).
- ²⁰ Fritz, P. J., and Hamrick, M. E., *Enzymologia Acta Biocatalytica*, **30**, 57 (1966).
- ²¹ Yap, H. H., and Cutkomp, L. K., *Life Sci.*, **9**, 1419 (1970).
- ²² Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J., *J. Biol. Chem.*, **193**, 265 (1951).

14 days after mating, and in those that became pregnant the interval between the stimulus and delivery of the litters was found to be 17–20 days. A second group of rabbits (group 2) was injected at approximately 21 days after mating, the interval between stimulus and delivery being 9–13 days. Samples of venous blood were obtained from each rabbit at the beginning of the experiment and at 10, 20 and 30 days after the injection of red cells. The sera were tested for haemagglutinating anti-Hg^A antibodies using papainized Hg^A-positive red cells.

Table 1 Anti-Hg^A Response in Rabbits Injected 14 days after Mating

	Pregnant rabbits 20 days after stimulus	30 days after stimulus	Non-pregnant rabbits 20 days after stimulus	30 days after stimulus
Anti-Hg ^A present	10 (55.6%)	10 (55.6%)	7 (41.2%)	9 (53.0%)
Anti-Hg ^A absent	8	8	10	8
Totals	18	18	17	17

Tables 1 and 2 give the results of testing for antibodies in the two groups of rabbits, which are divided into those that became pregnant as a result of the mating and those that did not, the latter acting as controls. (The overall incidences of anti-Hg^A in the control rabbits was 51% which is very close to the 56% incidence found in a group of non-mated control animals in a previous study⁶.) The incidence of positive tests for antibodies is given both for tests taken at the time when the pregnant rabbits reached the end of gestation and for tests taken 30 days after the stimulus. The latter are included because the pregnant state might have influenced not only the development of antibodies up to the point of delivery, but also their appearance in the post-gestational period. Previous experience⁶ had shown that if antibodies were to develop they had done so by 30 days after the stimulus.

Table 2 Anti-Hg^A Response in Rabbits Injected 21 Days after Mating

	Pregnant rabbits 10 days after stimulus	30 days after stimulus	Non-pregnant rabbits 10 days after stimulus	30 days after stimulus
Anti-Hg ^A present	7 (50%)	12 (85.7%)	11 (32.4%)	17 (50%)
Anti-Hg ^A absent	7	2	23	17
Totals	14	14	34	34

In group 1, samples taken 20 days after the stimulus, that is, at a time when pregnant rabbits reached the end of gestation, showed that 41.2% of non-pregnant rabbits had antibodies, while these were detected in 55.6% of pregnant animals. The difference is not significant. During the subsequent 10 days, two further non-pregnant rabbits developed antibodies, the final incidences in the two groups becoming almost identical. In group 2, tests or samples taken 10 days after the stimulus, that is, at a time when the pregnant rabbits were at the end of their gestation period, showed that 32.4% of non-pregnant rabbits had antibodies, compared with 50% of pregnant animals. During the subsequent 20 days a further eleven developed antibodies, the final incidences being 50% for non-pregnant and 85.7% for pregnant animals.

There is thus no evidence that during pregnancy the rabbits showed any diminished unresponsiveness to Hg^A-positive red cells. The peak titres of antibody reached in responding rabbits varied from 1:1 to 1:256, and there was no difference in the distribution of titres between pregnant and non-pregnant rabbits. There is a suggestion of an increased responsiveness

Effect of Pregnancy on the Isoantibody Response in Rabbits

WHEN an Rh-negative woman becomes pregnant for the first time with an Rh-positive foetus, the immune response to Rh-positive foetal red cells entering the maternal circulation is usually manifested by the appearance of antibodies after delivery or by priming without detectable antibodies, which only appear during a subsequent Rh-positive pregnancy¹. The appearance of antibodies during the first pregnancy, in the absence of any known prior stimulus with Rh-positive blood, is not as rare as was formerly thought. A minimum of 1% of such women can be found to have antibodies at the end of the first pregnancy, and they appear during the later months of ABO-compatible pregnancies².

Although the passage of foetal cells into the maternal circulation is most frequent and intense as the result of labour, it also occurs during pregnancy particularly during the third trimester². The relative infrequency of antibodies at the end of the first pregnancy and the frequent delay in their first appearance after delivery has led to the suggestion that women possess a diminished immune responsiveness to foetal red cell antigens during pregnancy^{3,4}, but the only direct experimental evidence is that of Merritt and Galton⁵, who found that the antibody response of three different species of experimental animal to sheep red blood cells, bacterial endotoxin and bovine γ -globulin respectively, was the same in pregnant as in non-pregnant animals.

To obtain further information on this point, but using an iso-antigenic system, random-bred Californian virgin female Hg^A-negative rabbits were mated and injected intravenously with 2×10^9 Hg^{AF} red cells at varying intervals after mating. One blood donor was used throughout the experiment. In an initial group of rabbits (group 1) the injections were given

in group 2 pregnant rabbits, but the difference from controls is not significant. The litters included both Hg^A-positive and Hg^A-negative rabbits, and although it was not possible to blood group all the newborn animals, analysis of those that were grouped showed that whether or not the mother developed antibodies was unrelated to the blood groups of the litter. Spontaneous immunization of Hg^A-negative rabbits by their Hg^A-positive foetuses has not been demonstrated⁷, and in the present experiment, animals carrying Hg^A-positive foetuses appeared no more tolerant of the injected Hg^A-positive red cells than did those carrying only Hg^A-negative foetuses.

Evidence for a depressed immune response during pregnancy comes largely from experimental allograft studies. It has been found that depression of allograft rejection during pregnancy may be (a) non-specific, that is, random grafts are better tolerated by pregnant than non-pregnant animals⁸⁻¹⁰, or (b) specific in that allografts possessing paternal antigens are better tolerated than other allografts^{9,10}.

It has been suggested^{10,11} that specific unresponsiveness may be related to the development in the mother of enhancing antibodies, and Currie was able to transfer such unresponsiveness to a paternal strain tumour with serum from interstrain pregnant mice. The ability to develop certain types of antibodies to paternal antigens may thus be an important event in pregnancy. Anti-leucocyte antibodies occur frequently in human pregnancy and are not uncommon even in a first pregnancy¹². This study shows that iso-antigens on red cells of the rabbit are capable of producing a normal antibody response in pregnant females.

We thank Professor Carl Cohen for typing rabbits and supplying antisera.

J. C. WOODROW
C. J. ELSON
W. T. A. DONOHOE

Nuffield Unit of Medical Genetics,
Department of Medicine,
University of Liverpool

Received May 13; revised June 1, 1971.

- ¹ Woodrow, J. C., and Donohoe, W. T. A., *Brit. Med. J.*, **4**, 139 (1968).
- ² Woodrow, J. C., *Rh Immunization and its Prevention* (Munksgaard, Copenhagen, 1970).
- ³ Gorman, J. G., Freda, V. J., Pollack, W., and Robertson, J. G., *Bull. NY Acad. Med.*, **42**, 458 (1966).
- ⁴ Ascari, W. Q., Levine, P., and Pollack, W., *Brit. Med. J.*, **1**, 399 (1969).
- ⁵ Merritt, Katherine, and Galton, M., *Transplantation*, **7**, 562 (1969).
- ⁶ Elson, C. J., Woodrow, J. C., and Donohoe, W. T. A., *Vox Sang.* (in the press).
- ⁷ Anderson, J. R., *Brit. J. Haematol.*, **2**, 44 (1956).
- ⁸ Heslop, R. W., Krohn, P. L., and Sparrow, E. M., *J. Endocrinol.*, **10**, 325 (1954).
- ⁹ Anderson, J. M., *Nature*, **206**, 786 (1965).
- ¹⁰ Currie, G. A., in *Foetal Autonomy*, Ciba Foundation Symposium (edit. by Wolstenholme, G. E. W., and O'Connor, M.), 32 (Churchill, London, 1969).
- ¹¹ Kaliss, N., and Dagg, M. K., *Transplantation*, **2**, 416 (1964).
- ¹² Overweg, J., and Engelfriet, C. P., *Vox Sang.*, **16**, 97 (1969).

Brain Catecholamines and Human Sleep

We have previously noted what seemed to be marked discrepancies between human and other mammalian species in the association of serotonin and sleep¹⁻³. Furthermore, L-dopa, the immediate precursor of dopamine and noradrenaline, by producing a decrease in human rapid eye movement (REM)

sleep⁴ (an opposite effect from the predictions based on animal studies⁵), indicated that there might also be differences between human and other mammalian species in the adrenergic system. To further determine the role of the adrenergic nervous system in human sleep we studied the effect of depleting brain catecholamines by administration of alpha-methyl-d-1-paratyrosine (AMPT) and alpha-methyl-P-phenylalanine (AMPH), which are potent and specific inhibitors of catecholamine synthesis^{6,7}.

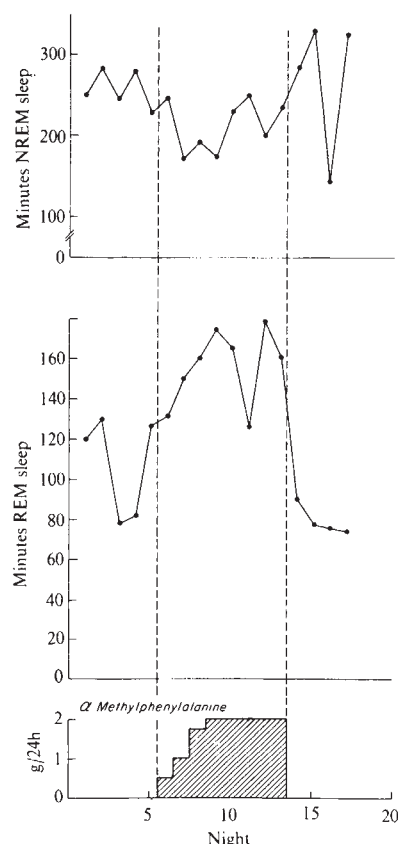


Fig. 1 Changes in rapid eye movement (REM) and non-rapid eye movement sleep (NREM) in patient (S. M., 42♀) taking alphanethylphenylalanine.

Five men and three women (two with essential hypertension, four with pheochromocytoma, one with Huntington's chorea and one with dystonia musculorum deformans) were studied by continuous all-night recording of electroencephalogram, electroculogram and electromyogram. Drugs were administered for therapeutic purposes. Each patient slept in his regular hospital room and, except for J. P. and B. F., who were allowed to sleep as late in the morning as they wished, was left undisturbed until awakened after 8 h of possible sleep. Previous medications were discontinued for at least 10 days before the beginning of the study. A constant number of identical capsules containing either placebo or active drug were administered four times each day with a maximum dose of 2-3 g per 24 h period. The protocol consisted of consecutive periods of pre-drug baseline placebo and active drug and post-drug placebo. Patients received no other drugs, and daytime naps were prohibited. Records were evaluated by standard criteria⁸ by a scorer unaware of the patient's drug status. The first 2 nights were used to adapt the patients to the slightly unnatural sleeping conditions and were not included in the results. Sleep records during the first 5 nights after the

daily drug dose exceeded 0.75 g were analysed. During pre-drug and post-drug placebo periods, data from 3–14 successive nights were examined.

In all eight patients, AMPT or AMPHe administration was associated with an elevation in total REM sleep (Table 1), ranging from 3–52%. This increase seemed to be related to an increase in the length of each REM period, because the total number of REM cycles was not affected. Three patients on AMPT had decreases in REM sleep after five nights, although the two other patients maintained the REM increase for the full period of drug administration (eight nights) as did the patients on AMPHe (Fig. 1). There was no change in non-rapid eye movement (NREM) sleep. When AMPT was dis-

available and in three further subjects who were receiving similar dosages of AMPT (Table 1). Lumbar punctures were performed in standard conditions⁹ during the pre-drug period and at the time of maximum AMPT dose and assays were carried out by methods previously described^{10,11}. The principal catabolite of dopamine (HVA), was reduced by 59%. The principal catabolite of serotonin 5-HIAA was increased in four of the five cases; however, these changes are not statistically significant. Because cerebrospinal fluid concentrations of monoamine catabolites seem to reflect central nervous system monoamine concentrations¹², these observations suggest that the AMPT in this study did produce a substantial reduction in brain catecholamines. The effect of AMPHe is pre-

Table 1 Effect of Alpha-Methylparatyrosine and Alpha-Methylphenylalanine on Duration of REM, NREM Sleep Time and Monoamines Metabolites

Patient	Total min REM			Total min NREM			Drug dosage in g/24 h	Pre-drug HVA	Placebo 5HIAA	AMPT HVA	Treat- ment 5HIAA	
	Pre-drug placebo	Drug treatment	Post-drug placebo	Pre-drug placebo	Drug treatment	Post-drug placebo						
AMPT	A. N.	65*	85	63	344	350	354	0.75-1.75	62	46	18	57
		(5)	(5)	(4)	(5)	(5)	(4)					
	B. R.	66	85	65	303	310	281	0.75-2.0	—	—	—	—
		(4)	(5)	(14)	(4)	(5)	(14)					
	R. L.	108	111	48	300	326	276	2.0-3.0	—	—	—	—
		(4)	(5)	(4)	(4)	(5)	(4)					
	J. P.	102	155	77	354	348	280	2.0	—	—	—	—
		(4)	(5)	(5)	(4)	(5)	(5)					
	H. D.	119	137	75	345	329	289	0.75-2.0	24	25	15	34
		(6)	(5)	(10)	(6)	(5)	(10)					
B. S.	—	—	—	—	—	—	0.75-2.0	14	28	≤4	31	
L. D.	—	—	—	—	—	—	0.75-2.0	20	34	16	27	
M. R.	—	—	—	—	—	—	0.75-1.75	18	50	≤4	58	
Mean ± s.e.	92±11	115±14 ⁺	66±50 ⁺⁺⁺	328±12	333±70	296±15		27±9	37±5	11±3 ⁺	41±7	
AMPHe	B. F.	106	142	—	348	400	—	2.0				
		(5)	(5)		(5)	(5)						
	S. M.	103	150	75	260	214	327	1.0-2.0				
		(5)	(5)	(4)	(8)	(5)	(4)					
	G. S.	91	100	—	285	246	—	2.0				
		(5)	(5)		(5)	(5)						

* Mean for number of nights given in parentheses. (+) $P < 0.05$ for difference from pre-AMPT treatment concentrations. (++) $P < 0.02$ for difference from AMPT treatment concentrations. (+++) $P < 0.05$ for difference from pre-AMPT treatment concentrations.

continued there was a decrease in both types of sleep. The NREM decrease (20–80%) from pre-AMPT baseline was seen in three patients for only 1–2 nights and is not reflected in the post-AMPT NREM sleep mean (Table 1). There was a REM sleep post-AMPT decrease (3–77%) in the five patients which persisted for up to six nights. In one patient (Fig. 1), studied after discontinuation of AMPHe there was a 25% decrease in REM sleep accompanied by a 27% increase in NREM sleep. A Friedman's analysis of variance of the overall REM means for the three was significant ($P < 0.05$). There were no significant changes for NREM sleep. All patients reported increased daytime sleepiness on AMPT but not on AMPHe.

As a test of the effect of AMPT on cerebral monoamine metabolism, cerebrospinal fluid concentrations of homovanillic acid (HVA) and 5-hydroxyindoleacetic acid (5-HIAA) were determined in two patients from whom sleep studies were

sumably similar to that of AMPT, as it also is a potent inhibitor of tyrosine hydroxylase.

Our results indicate that central catecholamine-containing systems take part in the control of sleep in man. The finding of a substantial increase in REM sleep in the absence of a significant change in NREM suggests that these sleep states can be pharmacologically dissociated. These observations differ from the results observed in animal studies in which catecholamine synthesis was inhibited. There are no reports on the effects of AMPHe administration on animal sleep but AMPT had variable effects on REM sleep in rats^{13,14} and cats^{15,16}, and decreased REM sleep in monkeys^{17,18}. Conversely, administration of L-dopa to cats whose monoamine stores were previously depleted by reserpine, restored REM sleep⁵. In humans, L-dopa given with and without a peripheral decarboxylase inhibitor (MK 485) decreased REM sleep⁴:

in several patients, L-dopa at higher dosages also decreased NREM sleep.

In contrast, our previous studies with parachlorophenylalanine¹⁻³ and 5-hydroxytryptophan¹⁹ indicate that in humans serotonin is intimately associated with the production of REM sleep, although a series of animal experiments suggest serotonin is associated with NREM^{5,20-22}.

We do not understand these apparent discrepancies between animal and human data, but they may be a result of differences in drug doses, route and period of administration or other parameters of experimental design³. It seems that in the human brain, however, the adrenergic and serotonergic brain systems act in opposite directions in the production of REM sleep, with the catecholamines decreasing, and serotonin augmenting, REM sleep.

Drugs were supplied by Dr Elmer Alpert, Merck Sharp and Dohme Co., Pennsylvania.

RICHARD J. WYATT
THOMAS N. CHASE
DAVID J. KUPFER
JIMMY SCOTT
FREDERICK SNYDER

National Institute of Mental Health,
9000 Rockville Pike,
Bethesda,
Maryland 20014

ALBERT SJOERDSMA
KARL ENGELMAN

National Heart Institute,
9000 Rockville Pike,
Bethesda,
Maryland 20014

Received August 11; revised December 17, 1970.

- ¹ Wyatt, R. J., Engelman, K., Kupfer, D. J., Scott, J., Sjoerdsma, A., and Snyder, F., *Electroenceph. Clin. Neurophysiol.*, **27**, 529 (1969).
- ² Wyatt, R. J., Chase, T. N., Engelman, K., Kupfer, D. J., Scott, J., Sjoerdsma, A., and Snyder, F., *Psychophysiology* (in the press).
- ³ Lovenberg, W., Engelman, K., Carpenter, jun., W. T., Wyatt, R. J., and Gessa, G. L., *Ann. Internal. Med.*, 607 (1970).
- ⁴ Wyatt, R. J., Chase, T. N., Engelman, K., Scott, J., and Snyder, F., *Nature* (in the press).
- ⁵ Jouvet, M., *Science*, **162**, 32 (1969).
- ⁶ Engelman, K., Horowitz, D., Jequier, E., Sjoerdsma, A., *J. Clin. Invest.*, **47**, 577 (1968).
- ⁷ Torcliana, M. L., Porter, C. C., Stone, C. A., and Hanson, A. M., *Biochem. Pharmacol.* (in the press).
- ⁸ *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects* (edit. by Rechtschaffen, A., and Kales, A.) (US Government Printing Office, Washington DC, 1968).
- ⁹ Chase, T. N., Schnur, J., and Gordon, E. K., *Neuropharmacol.* (in the press).
- ¹⁰ Cerbode, F., and Bowers, M. B., *J. Neurochem.*, **15**, 1053 (1968).
- ¹¹ Ashcroft, G. W., and Sharman, D. F., *Brit. J. Pharmacol.*, **19**, 153 (1962).
- ¹² Guldberg, H. C., and Yates, C. M., *Brit. J. Pharmacol.*, **36**, 535 (1969).
- ¹³ Marantz, R., and Rechtschaffen, A., *Perception Motor Skills*, **25**, 805 (1967).
- ¹⁴ Hartmann, E., Bridwell, I., and Shildkraut, J., *Psychophysiology* (in the press).
- ¹⁵ King, C. K., and Jewett, R. E., *Psychophysiology*, **6**, 220 (1969).
- ¹⁶ Iskander, T., and Kailbling, R., *Psychophysiology*, **6**, 219 (1970).
- ¹⁷ Crowley, J. T., Pegram, G. V., Smith, E., and Lewis, D. F., *Primate Electrophysiology Particularly Related to Sleep (Rep. ARL-TR-69-5)* (Holloman Air Force Base, New Mexico, 1969).
- ¹⁸ Weitzman, E. D., McGregor, P., Moore, C., and Jacoby, J., *Life Sci.*, **8**, 751 (1969).
- ¹⁹ Wyatt, R. J., Engelman, K., Sjoerdsma, A., and Snyder, F., *Psychophysiology* (in the press).
- ²⁰ Weitzman, E. D., Rapport, M. S., McGregor, P., and Jacoby, J., *Science*, **160**, 1361 (1968).
- ²¹ Koella, W. P., Feldstein, A., and Czicman, J. S., *Electroenceph. Clin. Neurophysiol.*, **25**, 481 (1968).
- ²² Torda, C., *Brain Res.*, **6**, 375 (1967).

Induction of Mesodermal and Endodermal Structures and Primordial Germ Cells in *Triturus* Ectoderm by a Vegetalizing Factor from Chick Embryos

A VEGETALIZING protein isolated from 9 day old chicken embryos¹⁻³ and implanted into the blastocoel of an early *Triturus* gastrula⁴ induces the ectoderm of the host embryo to form mesodermal tissues (muscle, notochord, renal tubes). Larger quantities of this highly purified factor cause the invaginated endoderm to exovaginate and spread over the induced ectoderm at the end of gastrulation. This may result from a change of cell affinities in the induced ectoderm⁵. In the exovagination series voluminous mesodermal and endodermal structures have been shown to develop from gastrula ectoderm under the influence of the implanted substance (U. K.-B., unpublished).

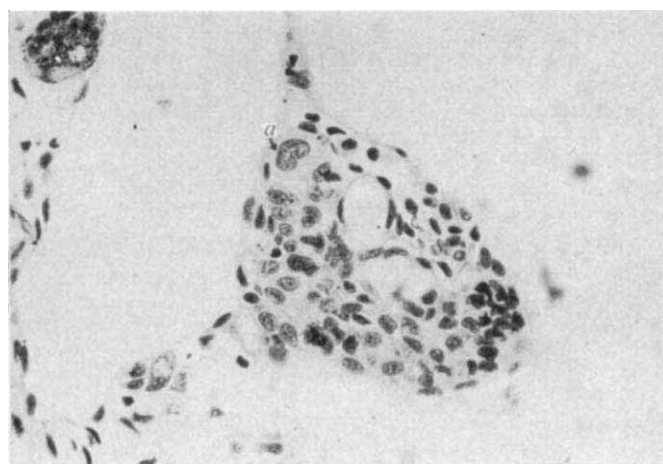


Fig. 1 Section through an induced gonad with an induced germ cell (a). Note the big cell with the typical potato-shaped nucleus. The explant was cultured for 41 days.

Effects of host embryos on the final outcome of the inductions are excluded if the inducer is tested on isolated gastrula ectoderm. To show the extent to which mesodermal and endodermal structures were induced by the factor, and to identify the induced endodermal organs, we used Holtfreter's sandwich technique⁶. Two pieces of gastrula ectoderm were sandwiched around a piece of the substance to be tested. Some explants were cultured at room temperature for 8-14 days, most of them for 3-7 weeks, to obtain good histological differentiation of endodermal organs. Most of the explants were spherical or cylindrical, compact and opaque and covered with endodermal epithelium.

As well as small unspecific neural structures, substantial but often poorly differentiated mesodermal and endodermal structures were induced in almost all cases. The induced mesodermal organs were notochord, muscle, renal tubes and sometimes mesenchym; the endodermal structures were mostly endodermal epithelium and intestine and, less frequently, liver, pancreas, stomach, oesophagus and pharynx. The relative frequency of all induced tissues varied between and within different experimental series, performed with various fractions of the vegetalizing factor. Though single explants consisted only of mesodermal tissues and others only of endodermal tissues and mesenteria, so far no fraction has induced mesodermal or endodermal organs exclusively. As well as these organotypic structures, in all series tissue complexes of large yolk-rich cells were found, which differed

clearly from typical undifferentiated ectoderm. If the time of action of the inducing factor is shortened by removal of the implant after 1–24 h or if the factor is diluted, differentiation of the induced tissues improves.

Besides these organs and tissues, in eleven out of 125 explants primordial germ cells were observed for the first time as a result of induction. The primordial germ cells were only found in explants which were cultured for at least 20 days. These especially large cells can be distinguished clearly from other cells by their histological appearance (Fig. 1). The explants in which primordial germ cells were found, contained also different mesodermal and endodermal tissues.

Both halves of the sandwich originate from gastrula ectoderm, so host germ cells, originating probably from lateral plate mesoderm^{7,8}, cannot have migrated into an induced germ ridge. The diagnosis of the primordial germ cells is based on their appearance and it has not been shown whether they are fertile. But the induction of these cells speaks strongly against the notion that the *keimbahn* is an absolute necessity for the development of germ cells in amphibians. Obviously germ cells can arise by regulatory mechanisms from somatic cells as do other cell types.

Probably only one factor induces the mesoderm–endoderm *anlage*, but the possibility that two factors are involved has not been excluded. All the observed differentiations including primordial germ cells certainly arise by secondary interactions.

URSULA KOCHER-BECKER

Heiligenberg-Institut,
Heiligenberg

HEINZ TIEDEMANN

Institut für Biochemie und Molekulare Biologie,
Freie Universität Berlin-Dahlem

Received March 22; revised June 14, 1971.

¹ Tiedemann, H., and Tiedemann, H., *Hoppe-Seylers Ztschr. Physiol. Chemie*, **306**, 7 (1956).

² Tiedemann, H., Born, J., Kocher-Becker, U., and Tiedemann, H., *Ztschr. Naturforsch.*, **20b**, 308 (1965).

³ Geithe, H. P., Tiedemann, H., and Tiedemann, H., *Biochim. Biophys. Acta*, **208**, 157 (1970).

⁴ Mangold, O., *Arch. Mikrosk. Anat.*, **100**, 198 (1923).

⁵ Kocher-Becker, U., Tiedemann, H., and Tiedemann, H., *Science*, **147**, 167 (1965).

⁶ Holtfreter, J., *Roux. Arch. Entw. Mech.*, **128**, 584 (1933).

⁷ Nieuwkoop, P. D., *Arch. Anat. Microsc. Morphol. Exp.*, **39**, 257 (1950).

⁸ Balinsky, B. I., *An Introduction to Embryology*, 45 (Saunders, London and Philadelphia, 1965).

Importance of Denitrification and Nitrate Reduction in Sediments to the Nitrogen Budgets of Lakes

DENITRIFICATION (biological conversion of nitrate and nitrite to nitrous oxide and molecular nitrogen) and nitrate reduction (assimilation of nitrate–nitrogen in microbial protoplasm) in lake sediments can play an important role in the nitrogen budgets of lakes. The hydrology of many lakes is such that ground water seepage through the floor constitutes a significant source of water income¹. In such cases, nitrate in the incoming ground water must be considered when a nitrogen budget is calculated for the lake system. An excellent example is the estimate for Lake Mendota, Wisconsin², in which nitrate in seepage water constituted nearly 36% of the nitrogen added to the lake.

Lake sediments, being highly reduced and rich in organic matter, are ideal for denitrification. Fermentation gases emanating from lake sediments contain large volumes of molecular nitrogen, presumably biogenic in origin³ and denitrifying microorganisms have been identified in marine and lake sediments (ref. 4 and unpublished results of R. L. Chen, D. R. K., J. G. Konrad, and A. J. Holding). Previous work in this laboratory⁵ showed that the nitrate and nitrite contents of Lake Mendota sediment profiles increased with depth below the sediment surface, indicating that nitrate was moving through the profile and being reduced. But the only report attempting to quantitate the extent of denitrification and nitrate reduction in lake sediments is that of Goering and Dugdale⁶ who in a laboratory incubation study with ¹⁵N-labelled nitrate, found that nitrate added to mud-water samples from Smith Lake, Alaska, disappeared within 6 days and that reduction of nitrate to ammonia was insignificant compared with reduction to molecular nitrogen. Other work^{7,8}, has indicated, however, that significant nitrate reduction and denitrification will occur concurrently in anoxic lake waters.

We have evaluated the extent of denitrification and assimilatory nitrate reduction in the surface sediments of Lake Mendota. In June 1970, sediment samples were collected with an Eckman dredge and treated in the boat either with 5 ml./l. sediment of ¹⁵N potassium nitrate solution (50.0 atom % ¹⁵N) equivalent to 2.1 mg per l. nitrate–nitrogen on a volume basis, or with 5 ml. of distilled water (control samples). The treated sediment was added to 1 l. plastic bottles which were returned immediately to the lake bottom at the sampling site. The experimental sites were under 7 or 14 m of water and their temperatures were 18° C and 12° C; the dissolved oxygen contents of the water just above the sediment surface of the two sites were 9 and 7 mg/l. The bottles were left uncapped throughout the experiment. At 4, 8, 12 and 16 days after initiation of the experiment, two of the control bottles and two containing the treated sediment were carefully retrieved for analysis. Investigations conducted by lowering and immediately raising sterilized nitrate–nitrogen-treated sediment showed that loss of nitrogen by wash-out was negligible (average 99.8% recovery). Similar investigations showed that less than 1% per day of nitrate–nitrogen was lost by diffusion into the lake water.

Interstitial water ammonium, nitrate and nitrite and sediment exchangeable ammonium, as well as sediment organic nitrogen (the nitrogen in the sediment residue after exhaustive 2 M KCl extraction), were determined⁵ and isotope ratioanalysis of interstitial and exchangeable ammonium–nitrogen and of sediment organic–nitrogen was conducted using a CEC 21–201 mass spectrometer by distillation and gas sample preparation methods^{9,10}. Physical and chemical properties of Lake Mendota sediments have been described^{5,11}.

We found that nitrate disappeared rapidly from the sediment samples; only trace levels of nitrate existed 4 days after initiation of the experiment and nitrite did not accumulate. Interstitial and exchangeable ammonium–nitrogen levels in the 14 m sediment samples (13 and 184 p.p.m. respectively) were about four times higher than those in the 7 m sediment and remained fairly constant throughout the experiment at both sites. Also, these values did not differ significantly between the control and the treated sediment at each site.

The ¹⁵N distribution data (Table 1) show that significant assimilatory nitrate reduction as well as denitrification was occurring in these sediments and that nitrogen turnover was quite rapid. Considerable incorporation of nitrate–nitrogen into microbial protoplasm occurred immediately on addition of nitrate, but in general the results reflect a microbial turnover of nitrogen coupled with sediment–water interchange. At both sites microbial immobilization was maximum at 4 days but was accompanied by considerable mineralization of protoplasmic nitrogen. About 37% of the ¹⁵N was recovered at 4 days, declining to about 20% at 16 days. The increase at 16

days was probably due to re-incorporation of soluble organic nitrogen compounds into microbial protoplasm, as soluble forms of nitrogen were not measured in our experiment. But this increase might be due to fixation of molecular ^{15}N produced during denitrification^{12,13}.

Table 1 Distribution of ^{15}N in Lake Mendota Sediment Samples at Various Times after Treatment with ^{15}N -labelled Potassium Nitrate

Time after initiation of experiment (days)	Recovery (%) of added ^{15}N in 7 m samples		Recovery (%) of added ^{15}N in 14 m samples	
	Organic-N	Ammonium-N	Organic-N	Ammonium-N
0.4	8.1	0.3	14.3	1.5
4	30.8	6.9	29.1	7.6
8	9.3	8.0	13.5	6.9
12	5.2	6.1	9.2	3.8
16	14.5	4.7	14.4	5.4

These data point out the fact that denitrification and assimilatory nitrate reduction occur concurrently in sediments and must be considered when evaluating the nitrogen budget of lakes receiving nitrate from ground water. Lee² has estimated that Lake Mendota receives about 36% of its nitrogen from ground water seepage. This amounts to about 77,000 kg nitrogen per annum for the lake. It is clear that if this figure were reduced by using the 4 day results (37% of the nitrate-nitrogen immobilized and 63% denitrified), the ground water portion of the Lake Mendota nitrogen budget could be revised accordingly. Such a revision shows that the nitrogen contribution would decrease by 49,000 kg due solely to consideration of sediment denitrification. This illustrates the point that the fate of ground water nitrate in lake sediments must be considered when nitrogen budgets of seepage lakes are calculated.

This research was supported by the College of Agricultural and Life Sciences, University of Wisconsin, and by a grant from the Federal Water Quality Administration. We thank Professor R. H. Burris for use of the mass spectrometer and Mrs K. Chen for assistance in the analyses.

D. R. KEENEY
R. L. CHEN
D. A. GRAETZ

Department of Soil Science,
University of Wisconsin,
Madison, Wisconsin 53706

Received May 6; revised June 4, 1971.

- Hutchinson, G. E., *A Treatise on Limnology*, 231 (Wiley, New York, 1957).
- Lee, G. F., *Eutrophication*, 17 (Univ. of Wisconsin Water Resources Center, Madison, 1970).
- Allgeier, R. J., Peterson, J., Juday, W. H., and Birge, E. A., *Intern. Review Hydrobiol.*, **26**, 444 (1932).
- Chamroux, M. S., *Ann. L'Inst. Pasteur*, **109**, 424 (1965).
- Konrad, J. G., Keeney, D. R., Chesters, G., and Chen, K., *J. Water Poll. Contr. Fed.*, **42**, 2094 (1970).
- Goering, J. J., and Dugdale, V. A., *Limnol. Oceanog.*, **11**, 113 (1966).
- Brezonik, P. L., and Lee, G. F., *Environ. Sci. Technol.*, **2**, 120 (1968).
- Goering, J. J., *Deep Sea Res.*, **15**, 157 (1968).
- Bremner, J. M., and Edwards, A. P., *Soil Sci. Soc. Amer. Proc.*, **29**, 504 (1965).
- Bremner, J. M., in *Methods of Soil Analysis* (edit. by Black, C. A.), 1256 (American Society of Agronomy, Madison, Wisconsin, 1965).
- Keeney, D. R., Konrad, J. G., and Chesters, G., *J. Water Poll. Contr. Fed.*, **42**, 411 (1970).
- Howard, D. L., Frea, J. L., Beeister, R. M., and Dugan, R. P., *Science*, **169**, 61 (1970).
- Brooks, R. H., Brezonik, P. L., Putman, H. D., and Keirn, M. A., *Limnol. Oceanog.* (in the press).

GENERAL

Uniform Distribution of Objects in a Homogeneous Field: Cities on a Plain

THE uniform distribution of objects in a homogenous field is familiar in all kinds of circumstances—atoms in liquids¹⁻³, fibrils in cells⁴, hairs on insects^{5,6} and mammals⁷, stomata on leaves⁸, trees in forests⁹⁻¹⁰ and cities on plains^{11,12}. In many of these circumstances the relative positions of the objects may be described by a radial distribution function. In what follows we shall show that such a function may be used to study the distribution of cities on a plain.

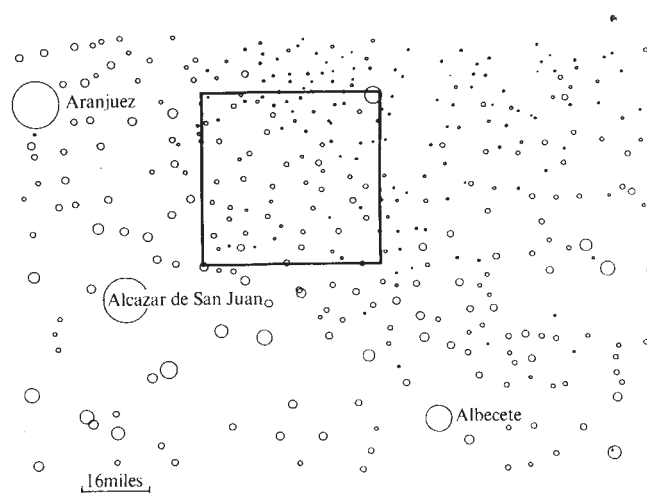


Fig. 1 A map of part of the Spanish Plateau lying south-east of Madrid, bounded by latitudes $0^{\circ} 53' \text{ W}$, $3^{\circ} 41' \text{ W}$ and longitudes $38^{\circ} 48' \text{ N}$, $40^{\circ} 16' \text{ N}$. The radial distribution function of the cities lying in the square has been computed as described in the text. The locations of the towns were taken from the US Army Map Service topographic sheets of the Iberian Peninsula, 1 : 250,000, dated 1944.

Fig. 1 is a map of part of the Spanish Plateau lying south-east of Madrid and bounded by latitudes $0^{\circ} 53' \text{ W}$, $3^{\circ} 41' \text{ W}$ and longitudes $38^{\circ} 48' \text{ N}$, $40^{\circ} 16' \text{ N}$. This area is homogeneous in climate, physiography, transportation, population density and economy. The towns are represented by circles whose radius is an approximate measure of physical extent¹³. We have chosen for analysis the 40 mile square centred at latitude $2^{\circ} 30' \text{ W}$ and longitude $39^{\circ} 47' \text{ N}$ which is especially homogeneous in town size and density.

The radial distribution function $g(R)$ is defined as the probability that a structure will be in a unit area a distance R away from an arbitrary structure, divided by the number probability density. It may be computed for the town lying in the square in Fig. 1 in the following way. Place each town in turn at the origin and surround it by a series of concentric circles spaced at a distance of $\Delta R = 1$ mile. Determine the area, $K_i(R)$ (lying wholly within the square in Fig. 1) of the annulus whose boundaries are at distances of $(R - \Delta R/2)$ and $(R + \Delta R/2)$ from town i . Let the number of towns lying within this annulus be $N_i(R)$. If N is the total number of towns in the square and K is the area of the square, then

$$g(R) = \frac{N(R)}{\rho K(R)}$$

where

$$N(R) = \sum_{i=1}^N N_i(R)$$

$$K(R) = \sum_{i=1}^N K_i(R)$$

$$\rho = \frac{N}{K}$$

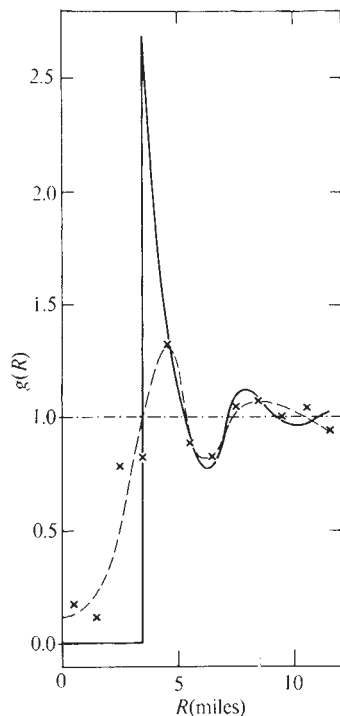


Fig. 2 The radial distribution function computed for the towns in the square in Fig. 1 (---), for a totally random distribution (— · —), and for a random distribution of hard-core disks with $\bar{\rho}=0.5$ and $d=3.46$ miles (—).

In Fig. 2 the computed distribution function is indicated by the dashed curve drawn through the crosses which represent the points at which the determination was made; the radial distribution function for a random distribution is depicted by the dashed-dot line in Fig. 2.

The diminished density near the origin and the subsequent damped oscillations in density are characteristic of distributions in which objects are fundamentally randomly dispersed, but are subjected to repulsive interactions which tend to diminish the probability of finding two objects close to each other. If the distribution is formed by a dynamic process which guarantees that the distribution is maximally random subject to the constraints of the interparticle interactions, then the theoretical concepts developed in the study of the statistical mechanics of equilibrium liquids may be applied.

An example is useful. Assume that the towns are randomly distributed subject only to the constraint that no two town centres may be closer than a distance, d . Similar models have been proposed as prototypes for spatial interaction and pattern formation in biological systems^{6,7,9}. Because the resulting distribution also arises in the hard-core model of classical liquids at equilibrium^{2,3}, the radial distribution

functions for such distributions have been intensively studied. The results are tabulated³ for different relative packing densities, $\bar{\rho}$, where $\bar{\rho}$ is the number density divided by the maximum close packed hard-core density, and is given by

$$\bar{\rho} = \frac{\sqrt{3}}{2} \rho d^2$$

The best agreement is given when $\bar{\rho}=0.5$, which corresponds to a hard-core town diameter of $d=3.46$ miles (solid line in Fig. 2). Although there is qualitative agreement, the hard-core interaction is too stringent to give detailed quantitative agreement with the experimental results. A more realistic interaction between towns would be graded, in which the probability of finding two towns close together would be low, but would continuously increase as the distance between the towns increased. This is what would be expected from the geographic theory of central places^{11,12}. Lawrence, in a study of hair placement in *Oncopeltus*, has hypothesized the existence of a very similar repulsive interaction which leads to the uniform distribution of hairs⁶.

No matter what system is considered, this hypothesized repulsive field plays much the same role as the intermolecular potential in the theory of liquids. If the distribution of structures is maximally random subject to the constraints imposed by this repulsive field, then the equations developed in statistical mechanics may be applied. It is possible to compute theoretical radial distribution curves for a given hypothesized field (ref. 1, chapter 5) and also to determine, approximately, the interaction between structures if the radial distribution function is known (ref. 1, page 80). Future work may characterize the hypothesized repulsive fields and explore their physical, biological and social origins in those diverse disciplines where uniform distributions are observed.

L. G. thanks Professor H. C. Longuet-Higgins for his help and the US Public Health Service, National Institute of Health, for financial support. The data were assembled by G. Aspbury during tenure of a National Science Foundation grant.

LEON GLASS

Department of Theoretical Biology,
939 E 57th Street,
University of Chicago,
Chicago, Illinois 60637

WALDO R. TOBLER

Department of Geography,
University of Michigan,
Ann Arbor, Michigan 48104

Received March 29, 1971.

- ¹ Egelstaff, P. A., *An Introduction to the Liquid State* (Academic Press, London, 1967).
- ² Alder, B. J., Frankel, S. P., and Lewinson, V. A., *J. Chem. Phys.*, **23**, 417 (1955).
- ³ Chae, D. G., Ree, F. H., and Ree, T., *J. Chem. Phys.*, **50**, 1581 (1969).
- ⁴ Farrell, R. A., and Hart, R. W., *Bull. Math. Biophys.*, **31**, 727 (1969).
- ⁵ Wigglesworth, V. B., *J. Exp. Biol.*, **17**, 180 (1940).
- ⁶ Lawrence, P. A., *Dev. Biol.*, **19**, 12 (1969).
- ⁷ Claxton, J. H., *J. Theoret. Biol.*, **7**, 302 (1964).
- ⁸ Ursprung, H., in *Major Problems in Developmental Biology* (edit. by Locke, M.) (Academic Press, New York, 1966).
- ⁹ Pielou, E. C., *J. Ecol.*, **48**, 575 (1960).
- ¹⁰ Pielou, E. C., *An Introduction to Mathematical Ecology*, ch. 10 (Wiley Interscience, New York, 1969).
- ¹¹ Christaller, W., *Central Places in Southern Germany* (Prentice Hall, New Jersey, 1966).
- ¹² Berry, B., and Horton, F., *Geographic Perspectives on Urban Systems* (Prentice Hall, New Jersey, 1970).
- ¹³ Tobler, W. R., *Area*, **1**, 30 (1969).

BOOK REVIEWS

Oppenheimer in Dock

The Oppenheimer Case, Security on Trial. By Philip M. Stern. Pp. xiv+591. (Rupert Hart-Davis: London, May 1971.) £3.50.

"THIS began as a book about a security case," writes Mr Stern in his preface. "It ended as a book about the security system. The case is now history. The system is still with us." A baleful opening for anybody old enough to remember the witch-hunting McCarthy period in the United States and the accusing of J. Robert Oppenheimer, theoretical physicist, so-called father of the atomic bomb, of being a Soviet agent. And not encouraging for anybody young who gives thought to the secret dossier-collecting activities of security organizations today—an estimated 14 million dossiers totted up by the US Department of Defense, 42 million by an American private investigation agency. . . . Mr Stern describes them: "Permanent collections of secret information, many of which may contain misinformation that the 'subject' has no way of rebutting or correcting—because he is not allowed access to it."

In the case of Robert Oppenheimer there were 3,000 pages of secret material, from which large black notebooks were made and circulated in advance of the hearings not just to the counsel who was going to lead the prosecution, but also to the judges—a tribunal called the Gray Board, set up by the US Atomic Energy Commission. So, by a reversal of the American "blank pad" (for the judges) rule in trials at law, Oppenheimer had laid on him from the start the burden of proving his own innocence.

Oppenheimer and his counsel were not allowed to see what was in the notebooks; and "for reasons of security" his counsel was sent out of the court while certain extracts were dealt with—not that this stopped the subsequent reading out in open court by the prosecution counsel of some of those which Oppenheimer's counsel had been denied. So Oppenheimer had to prove his own innocence in the face of accusatory material known to the judges that was never revealed in full to his side. Furthermore, several of the principal charges against him, including a crucial security charge, had been taken into account when he was given his Q-clearance seven years earlier; so he was in effect being "tried" twice over for the same thing. Furthermore, some of his "offences" were, at the time he was supposed to have committed them, not unlawful. . . . Security on trial indeed! If there was one thing security was *not*, it was on trial. Safe behind

closed doors, it totally evaded any critical, let alone judicial, investigation of its own processes and nature.

The history of the case is hair-raising. With an unforced combination of the historian's organization of fact and the novelist's interest in people, Mr Stern tells it brilliantly: each of the 500 pages simply has to be read. In the last fifty pages, when he specifically turns his attention to "security on trial", it loses a little of its coherence and grip, but that is a minor defect. He begins his story in 1936, when the young Oppenheimer, rich, cultivated, solitary-natured bachelor son of a Jewish East Coast family joined the Faculty of the University of California at Berkeley; and began in the "liberal" fashion of the times to associate with local Left wing organizations, near-Communists, Communists. With a sharp pertinence, Mr Stern remarks that when Oppenheimer was "tried" in 1954 these associations were set by his accusers and judges entirely against the current cold war situation of the early 1950s, and never judged against the historical background in which they occurred. To be fair to his accusers and judges, when reading about them now, one has to avoid the comparable trap of judging their behaviour entirely against the *détente* ambience of today, forgetting the paranoia and hysteria which was gripping the United States (not to mention the Soviet Union) then.

Throughout his career Oppenheimer showed an intellectual brilliance which permanently captivated his friends and his followers—and roused permanent hatred and alarm in the men he made his enemies. In his theoretical physics, though, he never made any fundamental discoveries: he was the quintessential solver of other people's problems rather than creator of his own. At first sight it seems odd that a scientist of this kind should succeed so astonishingly as the director of the scientists at the Los Alamos weapons laboratory who actually were making discoveries. Even odder that a man of his physical and temperamental type—essentially ectomorphic, fine-drawn, a creature of quick, nervous apprehensions—should be so masterful in the hurly-burly of power. And should continue to be so, later on, when from the eminence of the Princeton Institute for Advanced Study he engaged in controversy over whether the hydrogen bomb should be developed, and then whether it should be made the centrepiece of US defence—controversy in which he made an enemy of Edward Teller, father of the hydrogen bomb, and, more fatally, aroused the hatred of the US Strategic Air Command.

In fact the Gray Board's final judgment was that Oppenheimer "seemed to have had a high degree of discretion, reflecting an unusual ability to keep to himself vital secrets." Did most of them, and successive ratifiers of the verdict, ending with the Atomic Energy Commissioners and then the Government, bring him to his knees by revoking his security clearance really out of fear that at some future date he would give away secrets to the Russians? Or, Mr Stern asks: "Would Oppenheimer have been charged as a security risk if he had not held controversial opinions; if he had not advocated them with extraordinary persuasiveness, if he had not aroused (and on occasion sorely offended) persons in high places?"

A very good question. Given the answer is "No", it is disappointing not to find oneself whole-hearted about Oppenheimer as a man. It seems that in the last resort he lacked some kind of ballast—moral ballast, psychological ballast, possibly even intellectual ballast. Throughout the book examples abound to such an extent that it cannot be ignored. There was something very lightweight about the way he got himself mixed up with Communists before the Second World War. When he saw the first atomic device exploded he murmured a poetic quotation from the *Bhagavad-Gita*. And after having been so appalled by the destructive possibilities of the hydrogen bomb that he wanted America to forswear working on it and to get the mutual consent of Russia to do the same (what a hope!)—after all that he said, when Teller triumphantly showed it could be made, that it was "so technically sweet that you could not argue about it". So technically sweet. . . . There is a point at which lightness of weight, absence of ballast, can become faintly sickening. Yet it must be stressed that men who knew him, men such as I. I. Rabi and Hans Bethe, possessors of great ballast themselves, could reiterate years later their admiration and respect for him as a whole man. Mr Stern's attitude towards him, I may say, is impeccably detached and fair.

However, nobody's reading of Oppenheimer's character solves the enigma which overhangs the crucial security charge against him. Sometime in January or February 1943, Oppenheimer and his friend Haakon Chevalier were making martinis in Oppenheimer's kitchen. "There is no dispute over the fact," Mr Stern records, "that Chevalier reported that Eltenton said he had means of getting technical information to the Russians. It is also agreed that Oppenheimer made it immediately clear

that he would have no part in such an affair." The astonishing, inexplicable thing is that Oppenheimer later gave different accounts of the incident; and that in particular he gave an egregiously untruthful account to a security agent called Boris T. Pash, in which he refused to disclose Chevalier's name while saying that several other men were involved—a cock-and-bull story that was much more injurious to himself and Chevalier than the truth, and which at his "trial" he allowed to be called "a tissue of lies". It was an incredible thing to do; and when asked why he did it, he replied: "Because I was an idiot." It did him critical harm.

What could be the explanation? Nobody among his friends or his enemies has found a credible answer by looking at the character of Oppenheimer. Is it to be found, I propose, by looking instead at the character of the incident? Is that kind of interrogation liable to induce that kind of aberration? Oppenheimer's egregious performance *vis-à-vis* Colonel Pash was recorded through a hidden microphone and it makes disturbing reading—because of the peculiar intimacy existing between interrogator and interrogated. A fulsome, ingratiating duplicity on the interrogator's side; a fulsome, yielding over-accord on the interrogated's—"God bless you," he says at the end of it. One wants to turn one's head away . . . But that would mean turning one's head away from a very serious matter, where security really on trial could well be required to prove its case. The matter at issue is not the unnaturalness of the intimacy: it is the trustworthiness of the situation—its trustworthiness not to throw a man off balance, not to produce, while he is in it, almost a personality change. Is the solution to the Oppenheimer enigma not that he was an idiot, but that such interrogations are liable to bring idiocies to a man's lips? Really on trial, security is not invulnerable.

WILLIAM COOPER

Early Trains

Civil Engineering: Railways. By Bryan Morgan. Pp. xvi+176+45 Photographs. (Longman: London, April 1971.) £2.96.

THE admirable series of studies of industrial archaeology edited by Mr. L. T. C. Rolt has now reached railways. Wisely, he has decided that the subject is too big to be comprehended in a single volume, such as the one he wrote himself on *Navigable Waterways*, and that it must be treated in two. This book is therefore complemented by another on mechanical engineering, written by Mr J. H. Snell.

Mr Morgan's study is comprehensive, clearly arranged, and most pleasantly written. Within the modest limits en-

forced on him he packs a great deal of information and lively comment. He rightly stresses the importance of the early years of railway building, before the locomotive appeared and during its primitive, fitful development. He pays proper attention to the track, commenting with a severity that is well justified on the treatment usually accorded to "these unglamorous lengths of scrap iron" in museums, and taking trouble to explain the reasons that lay behind the adoption of different forms of rail, sleeper, and ballast. His illustrations and diagrams are a well integrated adjunct to his text.

The book has some faults. Its proportions may be criticized here and there. Scotland receives on the whole scanty attention. Of more than a hundred structures mentioned in the gazetteer at the end of the book, only four lie north of the border. Perhaps more serious, more than four-fifths of the text is devoted to the period before 1865. That was indeed the heroic age of railway building in this country, but notable work has been performed since, which is passed over here: for example, the extraordinary series of bridges, in Staffordshire blue brick, erected over the Midland Railway's line south of Kettering in the 1880s and 1890s, where the engineers were treating the bricks in an almost plastic way that seems to anticipate the handling of reinforced concrete in the twentieth century. And by way of a final criticism it must be said that there are rather more small mistakes here than there ought to be: obvious misprints uncorrected, names mis-spelt and places wrongly cited—Ashburton for Ashton, for example. It was not the North Midland Company that tried to reach Manchester (page 117); nor was Tite the architect of the station at Shrewsbury (page ix).

These small blemishes can be rectified when the book goes into a second edition. It deserves to be kept in print for a long time as an excellent introduction to its subject, which will not lose its value. Mr Morgan is generous in searching out distinction among his engineers and architects, whether in pioneering forms and techniques or in aesthetic sensibility, and he communicates his enthusiasm so attractively that he will inspire his readers to follow him in the direction of inquiry he has laid down. His book achieves its purpose very well indeed.

JACK SIMMONS

Röntgen Revealed

The Life of Wilhelm Conrad Röntgen. By W. R. Niske. Pp. xi+355. (University of Arizona: Tucson and Arizona April 19, 1971.) \$8.50.

WHEN in 1895 Röntgen made his great discovery he was already fifty years old,

and the rest of his life was in a sense an anticlimax; it seems that he genuinely hated fame and publicity, and regretted his move from Würzburg, where he had been happy, to Munich, where he was not. After his three papers describing the phenomena of the X-rays, he played no important part in the practical development of the apparatus or in the explanation of the phenomena; in this he reminds one of Volta, who made his discovery at much the same age. It is clear from this biography that Röntgen's discovery, while unexpected, was in no sense an accident. He had been carefully trained under A. E. E. Kundt to make very accurate determinations of physical quantities; first investigating the specific heats of gases, and then studying electromagnetism, and the physical properties of crystals. His academic position—he had held a chair at Giessen before being called to Würzburg to follow Kohlrausch—and his reputation for sober accuracy ensured that his preliminary paper on X-rays was taken seriously.

The study of cathode rays was an obvious field for investigation; mercury pumps were available to give very low pressures, and it seems that others had noticed fluorescence near cathode ray tubes and fogging of photographic plates, but had not seriously investigated the matter. In November and December 1895, Röntgen investigated the X-rays in an extremely capable manner, proving that the effects could not be due to cathode rays, and taking a number of X-ray photographs. He submitted his preliminary paper on December 28, 1895; it was published with a rapidity that moderns might envy, for the offprints were ready to be sent out on new year's day 1896. Röntgen was transformed from a man with a high reputation within his field into a celebrity; and was in 1901 awarded the first Nobel Prize for Physics.

Reading this sympathetic biography, one feels that Röntgen had somehow got out of his depth; like Faraday, he had the "ability to improvise laboratory equipment, which enabled him to make observations that less skilled men were unable to do", but we find little evidence that like Faraday he was a solitary voyager in strange seas of thought. His orderly life, with the spring and summer vacations spent in Italy and Switzerland, does not have that appearance of complete devotion to the struggle with the mysteries of nature which characterizes those few scientists of the first rank. Röntgen's example can encourage us all, for he seems to have been a rather ordinary scientist who made an extraordinary discovery; and this biography is valuable in bringing before us such a man and casting light on his times. In spite of a ten-

dency to degenerate into travelogue, the book is readable and interesting; we encounter both the crusty and forbidding professor and the shy man at ease holidaying with his friends. We even find him in 1888 complaining of the declining number of students because for scientists "the outlook is gloomy".

D. M. KNIGHT

Illuminating Letters

The Born Einstein Letters. By M. Born, Translated by I. Born. Pp. xi+240+4 plates. (Macmillan: London and Basingstoke, May 1971.) £3.85.

THIS volume consists of 120 letters between Einstein and Max and Hedwig Born, the first written by Einstein in 1916 as the General Theory was breaking on the world, the last by Born in January 1955, as both men became involved in the events leading up to the Russell-Einstein Manifesto. Commentaries written by Born in 1965, when he was already in his eighties, fill in details, add later judgments, and provide a background to the men, and the occasional woman, who lived through the dramatic era of physics between the two world wars.

The principal scientific interest in the correspondence lies of course in the long debate between the two men, and most, although not quite all, of their long discussions can be followed here. Their first meeting was at Salzburg in 1909 when Einstein, justy thirty, giving his first big invited paper, declared "a profound change in our views of the nature and constitution of light" to be indispensable. Born himself finds it hard to believe that he did not correspond with Einstein when working on the latter's theory of specific heats three years later.

The letters touch on the anti-Einstein movement of the early 1920s in Germany and on the confrontation between Einstein and Lenard at Bad Nauheim in 1920, in some ways comparable with the Wilberforce-Huxley dispute at Oxford in 1860. (Interestingly enough, no full verbatim account of either seems to have survived.) They add to the material on the birth of quantum mechanics already available in the Einstein-Sommerfeld letters and in the *Letters on Wave Mechanics* by Einstein, Schroedinger, Planck and Lorentz published four years ago. And they carry the debate on quantum mechanics into the 1950s.

Yet their interest does not lie solely in the contemporary comments of two masters, made as the new physics evolved. There is also their parallel reaction to events in Germany, to the birth of nuclear weapons and to the increased responsibility of scientists for

the ways in which their work is utilized. While Einstein signed the momentous letter to Roosevelt which helped launch the American atomic effort, Born found the idea of nuclear weapons so distasteful that he avoided personal involvement and tried—unsuccessfully—to dissuade one of his workers in Edinburgh from joining the British effort: Klaus Fuchs. And when Born returned from Britain to Germany after the Second World War, Einstein noted that he was going back "to the land of the mass-murderers of our kinsmen".

There is a third interest in the correspondence, in some ways almost as important as the light cast on the argument between the standard-bearers for two opposing scientific views. Both Max Born and Albert Einstein reveal in these letters not only great intellects, but a warmth and humanity sometimes considered—rightly or wrongly—to be rare among scientists.

R. W. CLARK

Physics Performed

Physics: Demonstration Experiments. Edited by Harry F. Meiners. Vol. 1: Mechanics and Wave Motion. Pp. x+1-654. Vol. 2: Heat, Electricity and Magnetism, Optics, Atomic and Nuclear Physics. Pp. iv+655-1395+36. (Ronald: New York, 1970.) \$30 per set.

ONE of the opening chapters of the two volumes entitled *Physics, Demonstration Experiments* was written by the late Sir Lawrence Bragg. His association alone makes these two books not just another pair of books on physics demonstrations but an inspiration—an inspiration to all those who wish to make a physics course in school or at university really alive.

The two volumes are divided into a subject classification of, volume I: *Mechanics and Wave Motion*, and volume II: *Heat, Electricity and Magnetism, Optics, Atomic and Nuclear Physics*. A great bonus in each book is an additional chapter: in volume I the whole philosophy and art of the use of demonstrations to teach physics is discussed, while in volume II the additional chapter discusses the use of closed circuit television, overhead projectors, stroboscopic effects and other modern teaching devices which are now becoming commonplace in the lecture room.

Not since *Demonstration Experiments in Physics* by R. M. Sutton has there appeared a work which deals with the subject so well. Professor Harry F. Meiners, the editor, has gathered the cream of international physics demonstration techniques and methods into these two volumes. Contributions from the United States, Britain, Israel and Russia are all included, together with many others.

From this mine of information for teachers and lecturers, it would be impossible to summarize, or even give a list of the demonstrations in these volumes, which together add up to 1,400 pages. It is easier to say that from a simple demonstration of shadow projection to solid state physics or the application of a laser, all are included.

On reading these books the complexity of some of the apparatus is a little worrying until it is realized that apart from the description of the demonstration in the text, full constructional details and drawings, almost down to the last nut and bolt, are given in an appendix at the end of each volume. If commercial apparatus is used a list of suppliers or manufacturers is given. This publication was sponsored by the American Association of Physics Teachers, and being American, most of the suppliers are American. This need not deter the English reader as equivalent equipment is obtainable in this country.

These books, which have some 2,000 photographs as well as line drawings, would be excellent additions to any Science Library.

To conclude, Sir Lawrence Bragg's advice to a lecturer was to not just talk about science, but, if possible, to demonstrate it. These books show the way.

W. A. COATES

Anaesthetic Collection

International Encyclopedia of Pharmacology and Therapeutics. Section 8, Vol. 1: Local Anesthetics. Section edited by P. Lechat. Pp. ix+377. (Pergamon: Oxford and New York, April 1971.) £6; \$16.

THIS volume of the *International Encyclopedia of Pharmacology and Therapeutics* comprises ten chapters written by internationally known authors, two of whom are now unfortunately deceased. It amply justifies the title encyclopaedia, and covers in exceptional detail the various aspects of local anaesthetics, which range from historical development to current clinical application, and deals with structure-activity relations, mode of action, pharmacological properties and their modification, methods for comparison of activity and finally the toxicity and side effects of these substances.

Each chapter is very fully referenced and the individual authors are to be complimented on the fullness of their text which is amply illustrated with photographs, diagrams and tables, although the first chapter on the historical development of local anaesthesia tends toward over illustration. A strong emphasis is placed throughout the book

on historical aspects and early work, and in several chapters careful reading is required, as the general theme passes through a maze of previous ideas, experimentation and thought, before more recent concepts are expounded and further compared. However, even though the book is often an exposition of previous and current work, an attempt is made when appropriate to provide critical assessment.

The part local anaesthetics play in the temporary alleviation of pain during surgery is well known, and the continuing search for newer and better agents serves to illustrate the importance of these drugs and the fact that even those of more recent development suffer from certain clinical deficiencies, a point stressed in the final chapter. Much knowledge has, however, still to be gleaned and this textbook with its well documented account of past development, the deficiencies and properties of current drugs, and the experimental techniques used for the assessment of new drugs, should do much to stimulate further interest in the subject and provide an excellent comprehensive reference and reading text, not only for pharmacologists, both graduate and undergraduate, but anyone with an interest in local anaesthetics.

A. J. DUXBURY

Pesticides Around Us

Pesticides in the Environment. By Robert White-Stevens. Pp. xv+270. (Marcel Dekker: New York, June 1971.) \$23.50.

THE current widespread interest in all forms of environmental pollution by chemicals ensures an extensive potential audience for any work with the title *Pesticides in our Environment*. This is the first of five books to be published as parts of two volumes under this title. This first part of the full volume, and the part yet to be published, together deal with the more theoretical aspects of pesticidal chemicals, whereas the second volume will deal with the more practical problems of handling pesticides in the environment. The editor of the series, Dr R. H. White-Stevens, now heads the Bureau of Conservation and Environmental Science in New Jersey, and has had considerable experience of research in the agricultural chemical field.

The professed aim of the series, as stated by the editor, is to present in authentic objective terms, the properties, functions, utility and contributions of pesticidal chemicals to human welfare, and to provide data that will enable the optimum combination of efficacy, safety, reliability, economy and environmental improvement to be attained.

Unfortunately, it is impossible to judge accurately from this first book in the series whether these ambitious aims will be achieved when both volumes have been published.

This book comprises three sections all written by well qualified authors. They discuss the chemistry and biology of pesticides (R. L. Metcalf), the metabolism of insecticides and fungicides (T. R. Fukuto and J. J. Sims) and the metabolism of herbicides (J. E. Loeffler and J. van Overbeek). The chapters are well written and serve as an excellent, chemically oriented, introductory text and reference work on pesticides. However, they are written so as to appeal more to practising scientists than to the layman, at whom they also profess to be aimed. It is a great pity that the book lacks an index, because this greatly lessens its value as a rapid reference text. One other shortcoming is that, if the discussion is to be really pertinent to the title of the series, the relative persistence of pesticides should be discussed, but it is not; perhaps this omission will be rectified in the second volume of this series. However, these criticisms are minor and do not detract from any appreciation of the efficient and workman-like way in which the sections are written.

The standard of production is good and the layout pleasing and suitable for the subject matter, but the price will probably prove a considerable obstacle to many potential readers. There is no doubt that this book is a useful new addition to the literature on pesticides, providing information that may not be readily available elsewhere.

CLIVE A. EDWARDS

Feathered Friends

Man and Birds. By R. K. Murton. (The New Naturalist.) Pp. xx+363 +32 plates. (Collins: London, May 1971.) £2.50.

CONFUSINGLY, an earlier volume in this New Naturalist series, by E. M. Nicholson, appeared under the title of *Birds and Men*. However, while Nicholson in 1951 could point out how little was known about the precise relationships of birds and men, Murton can show how much has been discovered. It is salutary that less than one-sixth of the references cited here were available to Nicholson when he wrote his book twenty years ago, and many of these are merely old, unquantitative records showing that a particular problem is not new. To quote but one example, Nicholson does not mention the oyster-catcher, which is thought by some to have a serious effect on the cockle industry; twenty years later, after the Ministry of Agriculture,

Fisheries and Food has spent perhaps £100,000 investigating the matter, Murton can devote thirteen pages to the problem based on many published works, mostly dating from about 1957.

Man and Birds contains an amount of theoretical population biology which the amateur may find somewhat heavy going, particularly the second and third chapters (on ecological considerations and predators and prey). Notwithstanding this, there are few places where such matters have been so forcefully put. The fundamental concept of population ecology is probably that of density dependent mortality; at certain seasons of the year, animals, as a result of reproduction, are usually commoner than some resources that they need. A number (usually relatively large) of them must die as a result of competition for the resource. If man kills a proportion of the animals it does not follow that breeding numbers will be reduced; usually they are not because individuals which would otherwise have died in competition with their fellows are now able to survive in place of them. This point, together with the rapid rate of increase (or recovery), when birds are scarcer than usual, is still often overlooked by people who seek to control or to protect birds.

Some 180 pages of the book cover, by species, the birds which impinge on man's activities; the section is of great value, for the author clearly knows his literature. Two final chapters cover birds and industrial man (including fouling of buildings and bird strikes and aircraft), and wildlife management (future methods of control, such as tape recorded alarm calls and distasteful compounds).

This is a very valuable review which will be widely read. However, since it was written for a market which is, to some extent, "popular" it could have been more simply written, and fewer digressions into methodology (and subjects such as photoperiodic physiology) would have made it more concise. So vast is the scope that no two people would agree on what should be included, but I would have included more about man's affect on the birds. Even the most ardent bird-hater must appreciate that massive mortality of birds may give him an early warning of something that he is doing which may ultimately back-fire on himself; large losses of sea birds, particularly the auks, are potentially important from this aspect. Also there is relatively little about bird problems outside Europe. The small finch *Quelea* in Africa does perhaps many times more damage than any birds in Europe. Problems in North and South America and elsewhere on rice-eating birds have been studied in some detail as have those of birds in vineyards, but these are barely mentioned.

CHRISTOPHER PERRINS

CORRESPONDENCE

Freedom of Scepticism

SIR,—In his letter (*Nature*, 232, 355; 1971) Dr Higginson raises a number of points that are perhaps representative of a certain kind of contemporary view and which I shall endeavour to answer.

All of us are influenced by a respect for the authority of properly qualified scientists and by the beliefs of our colleagues. But it is a voluntary alignment and every man must have the right to re-examine sceptically the verities of conventional wisdom.

There are also many well-qualified scientists who set constraints on these freedoms and react unfavourably when the expression of scepticism generates heat rather than light. But public disclosures are the primary means by which we mutually regulate our affairs! Friction may be generated by this exchange. But I believe this comes about when the sceptic is not afforded the same courtesies as is the true believer, and Dr Higginson's letter is a case in point.

For instance, there exist well-documented critiques of most of the published evidence concerning smoking and lung cancer and other health matters—critiques, moreover, that appeared in respected and rigorously reviewed journals. In the past few months articles have appeared by Dr Yerushalmy, evaluating the data on the effect of smoking on the weight and growth of infants born to mothers who smoke (*Amer. J. Epidemiol.*, 93, 443; 1971), and carefully re-examining and rejecting the claim advanced by a number of prestigious (and unreviewed) US Government reports that smokers suffer from a larger incidence of morbidity than do nonsmokers (*J. Amer. Stat. Assoc.*, 66, 251; 1971). The trouble, however, is that Dr Higginson does not care to acquaint himself with the literature of dissent or even acknowledge its existence (although demanding that we document our objections—*again?*).

Or Dr Higginson speaks of the "widely explored area of cigarettes and lung cancer"; but is this impression of such a thorough exploration not created in part, too, because Dr Hammond and others have published over and over again their findings based on a single study? Dr Hammond has produced over a dozen published papers, besides innumerable talks and testimonials, all based on his confidential 1960 population survey. (His procedures had been felt to be inadequate by a number of statisticians who, incidentally, have published their criticisms all along.)

On the other hand, how often can one publish a critical review?

Again, Dr Higginson betrays his extreme displeasure by equating a courteous request by fellow investigators to meet and discuss the possibility of a public review of data, including assurances of all possible safeguards, with a "smear" and hints that the perpetrator of this dark deed somehow is motivated by "the grant reported to have been made recently to his university by certain industrial interests". (I am sure that if Dr Higginson will address a proper request to the administration of Washington University and specify which industrial interest he seeks to expose, this information will be made available to him.)

I understand that centuries ago a number of members of the faculty of Oxford University protested the establishment of laboratories for students. Being men of the highest repute and integrity, they saw no need for students to verify their claims about what was observed in their laboratories at the cost of great bother to them and needless expense to the university. But this took place many years ago; and even at Oxford the faculty has proved to be wrong occasionally.

Yours faithfully,

T. D. H. STERLING

*Department of Applied
Mathematics and Computer
Science,
Washington University,
St Louis, Missouri*

Weaned from Confusion

SIR,—“The cow suckles the calf” and “the calf sucks the cow” are descriptions of the one phenomenon, so clear that there should be no difficulty; but there is. During recent fairly extensive reading about this and related topics, I was sometimes so confused by the haphazard use of the two verbs that I was at a loss to know who had been doing what to whom.

Among this reading was a letter¹ written by Cowie and his colleagues and published by you in 1951, which must evidently be blamed for some of the confusion, since some authors have quoted it as their authority for writing “suckle” when they have meant “suck”. Others have muddled the words quite indiscriminately and still others have adopted standard usage. Is it too late to agree that the language of the Psalms, St Matthew, Shakespeare and Lamb is good enough and that we should all use

it to record our observations and express our thoughts as clearly as we can?

Yours faithfully,

F. R. SPRATLING

*School of Veterinary Medicine,
Cambridge*

¹ Cowie, A. T., Folley, S. J., Cross, B. A., Harris, G. W., Jacobsohn, Dora, and Richardson, K. C., *Nature*, 168, 421 (1951).

Alluvial Gold

SIR,—In the issue of *Nature* for July 23 (232, 214; 1971) there is a brief item on the proposed exploration for gold and copper by Rio Tinto Zinc in the Mawddach estuary and Snowdonia. I was particularly interested in the statement that RTZ plans to sift or dredge the sediment of the estuary for alluvial gold.

During my doctoral studies from 1961 to 1964, I had occasion to study in detail the sediment from the estuary proper, from the freshwater rivers running into it, and from the offshore area in the immediate vicinity. In particular, I examined the texture and the accessory minerals of the sediments. The results of this examination showed significant similarities in the sediments from the estuary and the offshore environments and significant differences in the sediments from these two areas and that from the freshwater rivers (Mawddach and Wnion). This led me to the conclusion that "... it is very probable that the offshore areas are the major source for the sediment in the estuary, at least as far up as Penmaenpool ...".¹ Put more simply, this means that most of the sand in the Mawddach estuary is brought in from the sea and is *not* derived through the erosion of the gold-bearing rocks near Bontddu.

Also, as part of the overall study, I examined the records of the Barmouth Harbour Trust in the National Library of Wales, Aberystwyth. These records showed that a European firm had been granted a concession to dredge for gold in the Mawddach estuary just prior to World War II, but that the firm had abandoned the project as uneconomic.

The success of any dredging projects for alluvial gold in the Mawddach estuary, or even the need for preliminary exploration, are therefore questionable.

Yours faithfully,

R. M. McMULLEN

*951 Pinewood Crescent,
Ottawa, Ontario K2B 5Y3*

¹ McMullen, R. M., thesis, Univ. Reading (1964).

Obituary

Sir Lawrence Bragg



Fig. 1 William Lawrence Bragg as a young graduate at Cambridge. (By courtesy of the Nobel Foundation.)

SIR LAWRENCE BRAGG, who died on July 1 aged 81, had the unique distinction of having himself created the science to which he devoted his life's work, and lived long enough to experience its revolutionary impact, first on inorganic chemistry and mineralogy, then on metallurgy, and finally on organic chemistry and biochemistry. At the same time X-ray analysis itself advanced from the ingenious puzzle-solving of the early pioneers to the almost blindfold automation of the present day.

Röntgen discovered X-rays when Bragg was five years old, and von Laue, Friedrich and Knipping demonstrated X-ray diffraction by crystals in the spring of 1912, when Bragg was taking his physics degree at Cambridge (Fig. 1). Von Laue, who had predicted the effect, also succeeded in developing a theory of scattering from a three dimensional lattice, but he assumed that the crystal of zincblende used for diffraction had a primitive cubic lattice, and therefore could not explain why spots corresponding to certain indices were present while others with closely related indices were absent. He speculated whether the diffraction effects had arisen from fluorescence excited in the atoms of the crystal by the "aperiodic pulses" of the primary radiation or whether five prominent wavelengths pre-existing in the primary radiation had been separated out to

produce the different diffraction spots.

Bragg's father, then Professor of Physics at Leeds, heard of the German work in the summer of 1912 and discussed it with his son, who set about re-interpreting the X-ray pictures of zincblende published by the Munich group. Within a few weeks he presented the correct answer to a meeting of the Cambridge Philosophical Society. The German group had noticed that spots which were round when their photographic plate was close to the crystal became elliptical as the plate was moved further away. By a remarkable feat of imaginative insight Bragg realized that such a focusing effect would arise if the X-rays were reflected by successive atomic planes (Fig. 2), and he reformulated von Laue's conditions for diffraction into what became known as Bragg's Law, which gives a direct relationship between the crystal structure and its diffraction

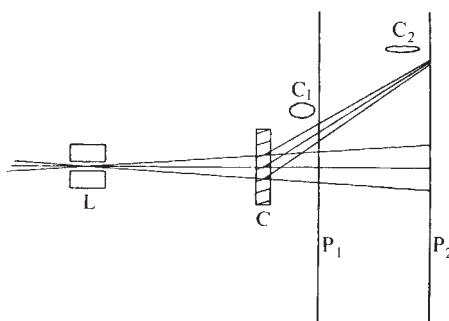


Fig. 2 Bragg's interpretation of the focusing of the diffracted spots observed by von Laue, Friedrich and Knipping which led to the formulation of Bragg's law: $n\lambda = 2d\sin\theta$. L, Lead screen; C, crystal; P_1P_2 , positions of photographic plate; C_1C_2 , cross sections of pencil of rays at P_1P_2 . (From *Proc. Camb. Phil. Soc.*, 17, No. 1, 43; 1913.)

pattern. Bragg then recognized that sets of parallel lattice planes would select from a continuous spectrum (or pulse, as he called it) those wavelengths which corresponded to integral multiples of the path difference between reflexions from successive atomic planes, so that each Laue spot would be made up of several harmonics of some selected wavelength. Finally, he demonstrated that the presence of spots with certain combinations of indices, and the absence of others, could be accounted for by assuming a face-centred rather than a primitive cubic lattice.

Why did this twenty-two year old student succeed in correctly interpreting the diffraction pattern predicted and discovered by an accomplished theoretician

eleven years his senior? Bragg himself modestly attributes it to a "concatenation of fortunate circumstances" but his paper soon convinces you that its success owed more to Bragg's astute powers of penetrating through the apparent complexities of physical mechanisms to their underlying simplicity.

This first paper was quickly followed by another, written in collaboration with his father, on their newly developed X-ray spectrometer, and a third, written by himself alone, solving the structure of common salt and showing how the Laue pictures of several other simple minerals could be indexed. There follows the structure of diamond, solved, as he relates, largely by his father, and the structures of fluorspar, zincblende, iron pyrites, calcite and dolomite solved by himself alone. Finally, on July 16, 1914, he communicated a paper on the structure of metallic copper. In view of this published record and the fact that for most of the relevant period the father was at Leeds and the son at Cambridge, it seems hardly believable that the scientific public tended to attribute most of the credit for these discoveries to his father, sometimes with the undertone that the son had cashed in on the father's success. The son must have suffered a great deal from these thoughtless judgments.

The outbreak of the First World War put an end to further research. Bragg joined the cavalry and was soon detailed to help in the development of sound ranging for the location of enemy batteries. At first he could not separate the report of the gun from the noise of the projectile travelling at supersonic speed. Then he suddenly spotted that the firing of the gun gave rise to a pressure wave, a "rush of air" as he called it, which he was able to detect by its momentary cooling of a hot wire. This simple device made an important contribution to victory.

He resumed his research when, aged twenty-nine, he was appointed Rutherford's successor as professor of physics at Manchester. His early research had replaced the concept of the inorganic molecule by the new one of a three dimensional array of close-packed and electrically compensating ions. The next step was a reformulation, on a structural basis, of the chemistry of the silicates. For example, diopside had been assigned the formula $\text{CaMg}(\text{SiO}_3)_2$, but B. E. Warren and Bragg discovered that there was no such thing as an SiO_3^{2-} ion. The structural units consisted of SiO_4 tetrahedra as in quartz; these shared corners so that they formed infinite chains bound to their neighbours

on all sides by Ca and Mg ions. The varying properties of the many different silicates were found to be a consequence of the ways in which SiO_4 tetrahedra can be linked, sharing corners, edges or faces, and forming long chains as in asbestos, sheets as in mica, or three dimensional networks as in the feldspars, so brilliantly solved by W. H. Taylor. Silicates make up the bulk of the Earth's crust, and Bragg was proud that he and his associates had explained their properties on the basis of their atomic structure. The results of the Manchester school also led Pauling to formulate general rules for the structures of inorganic crystals: atomic arrangements are governed by the ratios of the ionic radii and are stable only if all charges are locally compensated.

When the main structural principles of the silicates had been discovered, Bragg's interest turned to metals. He did not actually solve any metal structure himself, but became absorbed in the X-ray analysis of alloys by A. J. Bradley and others in the Manchester laboratory. His greatest papers of that period are those with E. J. Williams on order-disorder phenomena. For example, in brass at room temperature, copper occupies the corners and zinc the centres of a cubic lattice, but when heated they become more randomly distributed over the two sites until at a critical temperature all long range order ceases abruptly. Why does the approach to disorder suddenly become catastrophic? Bragg and Williams argued that "the decrease in order lessens the forces which tend to keep an atom in its 'right' place, since when disorder is nearly complete the criterion between 'right' and 'wrong' tends to vanish. The effects of increasing thermal agitation and lessening distinction between right and wrong so interact that on approaching a critical temperature the order drops rapidly to zero". Their physical explanation of the catastrophic approach to disorder in alloys later formed the basis for the interpretation of other cooperative phenomena in solids.

I first met Bragg in 1939 at the height of his metallic period, and he put me to work on the X-ray microscope, a neat optical device which he had just invented for the summation of two dimensional Fourier series. He was also keenly interested in my X-ray pictures of haemoglobin, but he did not actually work on the problem until the early fifties when he deduced the correct shape of the haemoglobin molecule from my data. Shortly afterwards he introduced the principle of minimum wavelength which limits the combination of signs in the Fourier transform of a molecule of given size. In the event, the limitations proved insufficient to solve the phase problem in haemoglobin, but the principle later helped D. L. D. Caspar and Rosalind Franklin to solve it in a radial projection of tobacco mosaic virus.

Bragg's last original paper concerned the determination of phase angles by the method of multiple isomorphous replacement with heavy atoms. In monoclinic crystals such as myoglobin or haemoglobin the location of the heavy atoms in the centrosymmetric projection was straightforward, but the determination of their third coordinate, which is not related to any symmetry element in the crystal, posed a difficult problem. Bragg devised an elegant method, based on a principle of physical optics, of deducing these coordinates from the changes in intensity produced by different isomorphous replacements in the same reflexions. As the correct values of the phase angles of several myoglobin reflexions emerged from his calculations he realized that the problem of protein structure, that seemingly hopeless venture which he had backed against all odds for the past eighteen years, could now be solved, and tears of emotion streamed down his face. Bragg's coordinates served as an independent

check for those which Kendrew and his collaborators deduced from Patterson syntheses. K. C. Holmes recently used his method to find the heavy atom coordinates in tobacco mosaic virus.

I have tried to summarize Bragg's many other contributions to science in a table which also shows the periods of his different activities in various places (Table 1). The sciences have developed so much faster than the arts. If we think of Bragg as an artist and compare him to, say, Giotto, it is as though he had himself invented three dimensional representation, and then lived through all the styles of European painting from the Renaissance to the present day, to be finally confronted by computer art.

So often men of genius were hellish to live with, but Bragg's creativity was sustained by a happy home life; typically one would find him tending his garden, with Lady Bragg, children and grandchildren somewhere in the background, and before getting down to business he would proudly demonstrate

Table 1 A Chronology of Bragg's Life and Progress

Year	Places and periods of activity		Important works not mentioned in the text, and books	
	Place	Period	Year	Work
1920	Manchester	Simple minerals	1921	Intensity of reflexion of X-rays by rock-salt (with R. W. James and C. H. Bosanquet)
			1922	Absolute intensities and the determination of structures with several independent parameters (with J. West)
			1924	Theory of double defraction in calcite
			1924	<i>X-rays and Crystal Structure</i> (with W. H. Bragg)
			1927	Isomorphous replacement (J. M. Cork)
1930	Manchester	Silicates	1929	Fourier projections
			1933	<i>The Crystalline State, A General Survey</i>
			1934	Order-disorder in metals (with E. J. Williams)
1940	NPL	Metals	1936	Structure factor graphs (with H. Lipson)
			1937	<i>Atomic Structure of Minerals</i>
			1939	X-ray microscope
			1940	Structure of cold-worked metal
1950	Cavendish Laboratory, Cambridge	Proteins	1942	Bubble model of plastic flow
			1945	Fly's eye
			1947	Dynamic model of a crystal structure (with J. F. Nye)
1960	Royal Institution	Teaching and popularization of science	1957	Determination of heavy atom parameters in proteins
			1965	<i>Crystal Structure of Minerals</i> (with G. F. Claringbull)
1970			1971	<i>The Development of X-ray Analysis</i>

his latest roses. To the present young generation Bragg was an avuncular figure who showed them that science can be fun. At the Royal Institution he instituted a series of physics lectures for sixth formers, making use of the institution's excellent experimental facilities to bring home the laws of optics, electricity and magnetism. These lectures were attended by 22,000 London school children each year. He also gave a remarkable series of television broadcasts on the properties of matter which brought him popular fame. He was intrigued when the greengrocer woman in Soho told him that he was "the spitting image of a man she saw on the telly last night" and modestly signed the bill for her to keep as a souvenir. Bragg's superb powers of combining simplicity with rigour, his enthusiasm, liveliness

and charm of manner, and his beautiful demonstrations all conspired to make him one of the best lecturers on science that ever lived. Yet few realized the formidable intellect that lay behind his genial façade or the decisive and revolutionary contributions to several fields of science for which he had been directly or indirectly responsible. His scientific papers still stand out as models of lucidity and conciseness. They generally contain little mathematics but are based on elegant physical and geometric arguments. His book on *The Crystalline State* once opened a new world to me and remains one of the best elementary introductions to X-ray crystallography.

Most X-ray crystallographers all over the world are "descended" from his or his father's pupils. He regarded them as one family and took the first steps in founding

the International Union of Crystallography soon after the end of the Second World War. J. C. Kendrew's and my own early attempts at protein crystallography could not have been sustained throughout the many lean years without his support, and our small group at the Cavendish Laboratory in Cambridge became the nucleus from which the present MRC Laboratory of Molecular Biology later developed. Several other schools, including H. Lipson's at Manchester and W. Cochran's at Edinburgh, are offshoots of his group at the Cavendish. D. C. Phillips's flourishing Molecular Biophysics Laboratory at Oxford stems from the small protein crystallography group which Bragg brought together at the Royal Institution in the 1950s. We all owe him a great debt.

M. F. PERUTZ

British Diary

Tuesday, September 7

Computers for Analysis and Control in Medical and Biological Research (three-day conference) Institution of Electrical Engineers, in association with the Institute of Electrical and Electronics Engineers, the Institute of Mathematics and its Applications, the Institution of Electronic and Radio Engineers, the Biological Engineering Society and the British Computer Society, at the University of Sheffield.

Displays (four-day conference) Institution of Electrical Engineers; and the Institution of Electronic and Radio Engineers, at the University of Technology, Loughborough.

Human Locomotor Engineering (four-day conference) Institution of Mechanical Engineers, at the University of Sussex, Falmer, Brighton, Sussex.

Thursday, September 9

Electrical Plasmas as Free-Atom Generators, and Reservoirs for Analytical Atomic-Absorption, Emission and Fluorescent Spectroscopy (2.30 p.m.) Professor V. A. Fassel, Society for Analytical Chemistry, Midlands Region, at Henry Wiggin and Co., Ltd, Holmer Road, Hereford.

Sunday, September 12

Cell Nucleus (Third Harden Conference, six days) Biochemical Society in collaboration with the Biological Council, at Wye College, Ashford, Kent.

Monday, September 13

Fundamentals of Software for Computer Engineers (vacation school, thirteen days) Institution of Electrical Engineers, in association with the Institution of Electronic and Radio Engineers, at the University of Manchester.

Surface Chemistry of Oxides (three-day discussion) Faraday Society, in the Chemistry Department, Brunel University, Uxbridge, Middlesex.

Reports and Publications

not included in the Monthly Books Supplement

Great Britain and Ireland

West of Scotland Agricultural College, Grassland Husbandry Department. *Experimental Records*. No. 22: Comparison of Productivity of Eight Varieties of Cocksfoot. By I. V. Hunt, J. Frame and R. D. Harkess. Pp. 16. No. 23: Varietal Productivity. Meadow Fescue, Series II. By I. V. Hunt, J. Frame and R. D. Harkess. Pp. 15. No. 24: Comparison of Productivity of Varieties of Tall Fescue. By J. Frame, R. D. Harkess and I. B. Hunt. Pp. 45. No. 25: Comparison of Productivity of Varieties of Timothy. By I. V. Hunt, J. Frame and R. D. Harkess. Pp. 38. (Auchincruive, Ayr: West of Scotland Agricultural College, 1971.) [26]
The Prospects for a European Security Conference. By Michael Palmer. (European Series, No. 18.) Pp. 107. (London: Chatham House and PEP, 1971.) £1. [36]
The Birds of Exmoor. By N. V. Allen. Pp. 644+14 plates. (Dulverton, Somerset: The Exmoor Press, 1971.) 50p. [36]
Department of Health and Social Security. Buildings for Mentally Handicapped People. Pp. 99. (London: HMSO, 1971.) £2. [36]

HOW TO BUY NATURE

Volumes start in January, March, May, July, September and November, but subscriptions may begin at any time.

The direct postal price per subscription is:

12 MONTHS* (52 issues per title)

	Surface Mail UK and worldwide	Airfreight U.S.A.	Canada
Nature (Friday)	£14	\$48	\$52
Nature + Nature Physical Science	£24	\$83	\$90
Nature + Nature New Biology	£24	\$83	\$90
All three editions	£29.50	\$108	\$116
Annual Index	£1	\$3	\$3

* Rates for shorter periods *pro rata* (minimum three months)

(Charge for delivery by air mail on application)

Editorial and Publishing Offices of NATURE

MACMILLAN JOURNALS LIMITED
4 LITTLE ESSEX STREET, LONDON WC2R 3LF
Telephone Number : 01-836 6633. Telegrams : Phusis London WC2R 3LF

711 NATIONAL PRESS BUILDING
WASHINGTON DC 20004
Telephone Number : 202-737 2355

Subscription Department
MACMILLAN JOURNALS LIMITED
BRUNEL ROAD, BASINGSTOKE, HANTS
Telephone Number : Basingstoke 5431

American display advertisements
NATURE SCIENTIFIC PUBLICATIONS INC
711 NATIONAL PRESS BUILDING
WASHINGTON DC 20004

All other advertisements
T. G. SCOTT & SON, LIMITED
1 CLEMENT'S INN, LONDON WC2A 2ED
Telephone : 01-242 6264/01-405 4743
Telegrams : Textualist London WC2A 2ED

Registered as a newspaper at the Post Office

Copyright © Macmillan Journals Limited, September 3 1971